

A study of consonant phonology in the Gāndhārī language
Insights into patterns of language variation and change

John Stuart Ryder

A dissertation submitted in fulfillment of requirements
for the Degree of Doctor of Philosophy

School of Languages and Culture
Faculty of Arts and Social Sciences

The University of Sydney

March 2024

This is to certify that to the best of my knowledge the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all assistance received in preparing this thesis and sources have been acknowledged.

.....

John Stuart Ryder

Abstract

The consonant phonology of the Gāndhārī language is described in detail in this thesis. Gāndhārī is a Middle Indo-Aryan (MIA) language which was spoken in ancient Gandhāra, encompassing areas of modern-day northern Pakistan and eastern Afghanistan. The use of Gāndhārī in the region is attested from the third century BCE up to around the fourth century CE, through sources which include major rock inscriptions and a significant number of Buddhist manuscripts. It is typically recorded in the Kharoṣṭhī script, which was developed from the Aramaic script, and does not accurately convey all features of the Indian sound system. While the Gāndhārī consonant phonology has been described in a number of existing publications, these are typically focused on individual inscriptions and manuscripts, and prepared in the wider context of Buddhist studies. This material has been collated into a systematically organised format, offering a comprehensive and accessible Gāndhārī language resource for use by Gāndhārī and linguistic scholars. This has facilitated linguistic analysis of the changes which occurred in Gāndhārī to the consonants and consonant clusters of the Old Indo-Aryan (OIA) ancestor languages.

An interesting feature of the Kharoṣṭhī inscriptions and manuscripts is the widespread use of variable orthographic forms to represent the consonant phonemes. These include historic and advanced forms as well as forms which appear to represent allophones arising in particular environments, suggesting the co-existence of the sounds represented by these variable forms in the spoken language. Combined with evident levels of uncertainty about the correct representation of the language by the scribes, the Kharoṣṭhī materials provide an invaluable record of patterns of language change in progress in a two thousand year old language. A system for categorising the variable phonemic forms is proposed, with the aim of developing a practical shorthand method for recording them, while also facilitating a simpler representation of these variations and analysis of the associated sound changes. Consideration is given to selecting the most appropriate phonemic representation of the Gāndhārī phonemes for use in the lemmas of words. Comparison of the Gāndhārī phonological record with the phonologies of other MIA languages, including Pali, offers insights into the processes of manner assimilation which occurred in the OIA consonant clusters, while aspects of the nature of Pali are reviewed in this context. Consideration is also given to sociolinguistic factors which may

have influenced the overall simplification which occurred in the grammar of the MIA languages.

Acknowledgements

It would not have been possible for me to initiate, let alone develop and complete this thesis without the generous and ongoing support, encouragement and advice from my supervisors, Dr. Mark Allon and Dr. Mark Post at the University of Sydney. Dr Allon has willingly shared his remarkable knowledge of the Gāndhārī language and Kharoṣṭhī script, and of the wide range of Gāndhārī source materials and publications. His generosity extended to accepting a semi-retired engineer as one of his students, initially for an honours thesis and subsequently for this dissertation. He clearly has a strong interest in advancing the knowledge of Gāndhārī, and has always been open and welcoming to my interest in studying the language from a linguistic perspective. I first met Dr. Post as a lecturer in linguistics at the University of New England and was immediately impressed by his deep knowledge and passion for linguistics. When I approached him at a later stage at the University of Sydney to act as Auxiliary supervisor for both my honours thesis and this dissertation, he was at once interested and generous in his support. Without the guidance and direction of both of my supervisors, I would not have been able to develop this thesis in the ways I have.

My interest in the OIA and MIA languages began with the study of Sanskrit, and I owe a great deal to Dr. McComas Taylor who teaches a wonderful online Sanskrit course at the Australian National University. Like many students I was inspired by his passionate interest in Sanskrit, which provided me with a strong foundation for an ongoing interest in the OIA and MIA languages. My thanks are also extended to the many earlier and contemporary Gāndhārī scholars from around the world, whose works I have been able to draw on in compiling this thesis. Although I have only been in contact with some of these scholars in person, at least through e-mail or on Zoom in these Covid times, I am particularly indebted to Dr. Richard Salomon, Dr. Timothy Lenz and Dr Joseph Marino of the University of Washington in Seattle in the USA for their support at different times. I would also like to thank Dr Yasuko Suzuki of the Kansai GaiDai University in Japan for kindly forwarding some of her papers, and sharing her knowledge of the Pali assimilation patterns.

My thanks are also extended to Ian McCrabb and Stephanie Majcher, fellow students of Gāndhārī at the University of Sydney during the early years of my studies, for sharing their knowledge of Gāndhārī and the digital tools they have developed.

Table of Contents

Abstract	iii
Acknowledgements	v
Maps and Tables	xv
Symbols	xix
Abbreviations	xxi
Explanation of Terms	xxiii
1. Introduction	1
1.1 Consonant phonology of the Gāndhārī language	1
1.2 An outline of previous scholarship	2
1.3 Aims of this thesis	4
1.4 Methodologies	6
1.4.1 Collation and summary of materials	6
1.4.2 Analysis of phonological change between OIA and Gāndhārī	7
1.4.3 Assessment and representation of variant Gāndhārī reflexes	8
1.4.4 The Gāndhārī record of language change in progress	8
1.4.5 Comparative development with other MIA languages	11
1.4.6 Manner assimilation of OIA consonant clusters in MIA	12
1.5 Overview of chapters	12
2. The Gāndhārī language	15
2.1 Overview	15
2.2 Linguistic environment	19
2.2.1 Historical background	19
2.2.2 The OIA languages	20
2.2.3 The MIA languages and Pali	21
2.2.4 Languages or dialects?	25
2.2.5 Source materials utilised by the Gāndhārī scribes	27
2.2.6 Dardic languages	28
2.3 A sociolinguistic perspective	29
3. Gāndhārī source materials	32
3.1 Inscriptions	32

3.2 Administrative documents	33
3.3 Buddhist manuscripts	33
3.4 Summary of source materials	36
3.5 Selection of source materials for analysis	36
3.6 Selected source materials	47
4. Sanskrit phonetics and phonology	49
4.1 The Sanskrit phonological system	49
4.2 The Sanskrit sandhi system	51
4.3 Sanskrit nasal consonants	53
4.4 Sanskrit consonant clusters	54
4.5 Sanskrit syllable structure	55
4.6 Sanskrit geminate consonants	57
5. The Kharoṣṭhī script and Gāndhārī phonological system	60
5.1 The use of syllabic modifiers in the Kharoṣṭhī script	63
5.2 Phonetic modification of intervocalic consonants	63
5.3 Consonant clusters in the Kharoṣṭhī script	66
5.3.1 Characters which record reflexes of original OIA clusters	67
5.3.2 Conjunct consonants recorded using diacritic marks	70
5.3.2.1 C{v, y, r} clusters	70
5.3.2.2 r-C clusters	72
5.3.2.3 Anusvāra	73
5.3.3 Conjunct consonants recorded using unique characters	74
5.3.4 Conjunct consonants recorded using combination characters	75
6. Patterns of sound change	78
6.1 Assimilation in consonant clusters	78
6.1.1 Perceptual and phonetic factors	79
6.1.2 Place assimilation	79
6.1.3 Manner assimilation	81
6.1.4 Synchronic and diachronic considerations	84
6.2 Lenition, deletion and fortition	84

7. Recovering the phonology of the Gāndhārī language	86
7.1 General methodology	86
7.2 Interpretation of orthographic intent	88
7.3 Representation of Gāndhārī phonemes in the lemma form of words	91
8. From Old Indo-Aryan to Gāndhārī: processes of phonological change	95
8.1 Introduction	95
8.1.1 The ordering of consonants in tabulations and charts	97
8.1.2 The recording of geminates which are reflexes of consonant clusters	98
8.2 Changes to uncombined OIA consonants in Gāndhārī	99
8.2.1 Stops	99
8.2.1.1 Word initial	99
8.2.1.2 Word medial	100
8.2.1.2.1 Velar stops	103
8.2.1.2.2 Palatal stops	110
8.2.1.2.3 Retroflex stops	113
8.2.1.2.4 Dental stops	114
8.2.1.2.5 Labial stops	120
8.2.1.2.6 Overview – intervocalic stops	124
8.2.1.2.6.1 Reflexes recorded using diacritic markers	126
8.2.1.2.6.2 Reflexes recorded using the graphemes of other phonemes	127
8.2.1.2.6.3 Summary and phonemic representation	131
8.2.1.3 Word final	132
8.2.2 Nasals	135
8.2.2.1 Word initial	135
8.2.2.2 Word medial	135
8.2.2.2.1 The labial nasal <i>m</i>	135
8.2.2.2.2 The dental and retroflex nasals <i>n</i> and <i>ṇ</i>	136
8.2.2.3 Word final	138
8.2.3 Fricatives	139
8.2.3.1 Word initial	139

8.2.3.2	Word medial	140
8.2.3.3	Word final	146
8.2.4	Semi-vowels	146
8.2.4.1	Word initial	146
8.2.4.2	Word medial	146
8.2.4.3	Word final	152
8.2.5	Scribal variations	152
8.2.5.1	Inclusion of diacritic markers to distinguish sound changes	152
8.2.5.2	Variable use of <i>n</i> / <i>ṇ</i>	153
8.2.5.3	Elision of unaspirated velar and palatal stops in palatal environments	153
8.2.5.4	Elision of <i>y</i> in palatal environments	153
8.2.5.5	Overview	154
8.2.6	Summary of changes – uncombined single consonants	154
8.2.6.1	General patterns of change	154
8.2.6.2	Gāndhārī phonological system	155
8.2.6.3	Comparison with Torwali phonological system	155
8.3	Changes to OIA consonant clusters in Gāndhārī	159
8.3.1	Stop plus clusters	159
8.3.1.1	Stop + stop clusters	159
8.3.1.2	Stop + nasal clusters	165
8.3.1.3	Stop + fricative clusters	170
8.3.1.4	Stop + semi-vowel clusters	172
8.3.1.4.1	Stop + <i>l</i> clusters	172
8.3.1.4.2	Stop + <i>v</i> clusters	175
8.3.1.4.3	Stop + <i>y</i> clusters	178
8.3.1.4.4	Stop + <i>r</i> clusters	184
8.3.2	Nasal plus clusters	190
8.3.2.1	Nasal + stop clusters	190
8.3.2.2	Nasal + nasal clusters	194
8.3.2.3	Nasal + fricative clusters	195

8.3.2.4	Nasal + semi-vowel clusters	197
8.3.2.4.1	Nasal + <i>l</i> clusters	197
8.3.2.4.2	Nasal + <i>v</i> clusters	197
8.3.2.4.3	Nasal + <i>y</i> clusters	199
8.3.2.4.4	Nasal + <i>r</i> clusters	202
8.3.3	Fricative plus clusters	204
8.3.3.1	Fricative + stop clusters	204
8.3.3.2	Fricative + nasal clusters	215
8.3.3.3	Fricative + fricative clusters	220
8.3.3.4	Fricative + semi-vowel clusters	221
8.3.3.4.1	Fricative + <i>l</i> clusters	221
8.3.3.4.2	Fricative + <i>v</i> clusters	221
8.3.3.4.3	Fricative + <i>y</i> clusters	225
8.3.3.4.4	Fricative + <i>r</i> clusters	229
8.3.4	Semi-vowel plus clusters	234
8.3.4.1	<i>l</i> + C clusters	234
8.3.4.1.1	<i>l</i> + stop clusters	234
8.3.4.1.2	<i>l</i> + nasal clusters	235
8.3.4.1.3	<i>l</i> + fricative clusters	236
8.3.4.1.4	<i>l</i> + semi-vowel clusters	236
8.3.4.1.5	Summary of <i>l</i> + C cluster developments	237
8.3.4.2	<i>v</i> + C clusters	237
8.3.4.2.1	<i>v</i> + stop clusters	237
8.3.4.2.2	<i>v</i> + nasal clusters	237
8.3.4.2.3	<i>v</i> + fricative clusters	238
8.3.4.2.4	<i>v</i> + semi-vowel clusters	238
8.3.4.2.5	Summary of <i>v</i> + C cluster developments	240
8.3.4.3	<i>y</i> + R clusters	241
8.3.4.4	<i>r</i> + C clusters	241
8.3.4.4.1	<i>r</i> + stop clusters	242
8.3.4.4.2	<i>r</i> + nasal clusters	246

8.3.4.4.3	<i>r</i> + fricative clusters	248
8.3.4.4.4	<i>r</i> + semi-vowel clusters	250
8.3.4.4.5	Summary of <i>r</i> + C cluster developments	254
8.3.5	Three consonant clusters	256
8.3.5.1	Stop + clusters	258
8.3.5.2	Nasal + clusters	259
8.3.5.3	Fricative + clusters	259
8.3.5.4	Semi-vowel + clusters	260
8.3.5.5	Overview	261
8.3.6	Summary of Gāndhārī reflexes of OIA consonant clusters	262
8.3.6.1	Word initial clusters	262
8.3.6.2	Word medial clusters	263
8.3.6.2.1	No change	265
8.3.6.2.2	Place assimilation only	265
8.3.6.2.3	Manner assimilation	265
8.3.6.2.4	Lenition / fortition	269
8.3.6.3	Three consonant clusters	269
8.3.6.4	Scribal variations	269
8.3.7	A record of language change in progress	272
8.4	Changes to Sanskrit syllable patterns in Gāndhārī	273
8.4.1	Changes to consonant clusters – geminate reflexes.....	273
8.4.2	Changes to single consonants	274
8.4.3	Changes to vowels and vowel sequences	274
8.4.4	Syllable patterns in Torwali	275
8.5	Retroflex consonants in Gāndhārī	276
8.6	Consonant sandhi phenomena in Gāndhārī	278
8.7	Later stage developments: third century CE inscriptions and manuscripts ...	279
9.	Comparative developments between Gāndhārī, MIA and Pali	281
9.1	Word initial consonant clusters	281
9.2	Word medial consonants and consonant clusters	282
9.2.1	Single consonants	282

9.2.2	Consonant clusters	283
9.2.2.1	Comparative development: Aśokan inscriptions	283
9.2.2.2	Comparative development: Gāndhārī and later stage Prakrits	285
9.2.3	Comparative overview: medial consonants and clusters	287
9.3	Review of MIA manner assimilation in medial clusters	290
9.3.1	Strength based accounts	290
9.3.2	A composite approach	292
9.3.3	Sanskrit gemination patterns	292
9.3.4	An analytical approach: articulatory characteristics and syllable patterns	294
9.3.4.1	Clusters having an increasing strength profile	294
9.3.4.2	Clusters having a decreasing strength profile	295
9.3.4.3	Summary	297
9.3.5	Sociolinguistic considerations	297
10.	Common types of sound change recorded in Gāndhārī consonants	301
10.1	Epenthesis	301
10.2	Palatalisation	304
10.3	Deaspiration	305
10.4	Aspiration	306
10.5	Retroflexion	306
10.6	Metathesis	306
10.7	Occlusion and loss of nasality in sibilant + labial consonant clusters	308
11.	Conclusions	309
11.1	Collation and analysis of the Gāndhārī phonological record	310
11.2	Categorisation of the variable phonemic record in a shorthand format	310
11.3	Changes to OIA single consonants in the Gāndhārī	311
11.3.1	Summary of distinctive phonological features of Gāndhārī	311
11.3.2	Interpretation of phonemes and phonemic representation	311
11.3.3	The Gāndhārī phonological system	313
11.4	Changes to OIA consonant clusters in Gāndhārī	313
11.4.1	Summary of distinctive features of consonant clusters in Gāndhārī	313

11.4.2 Manner assimilation in OIA consonant clusters.....	314
11.5 A record of language change in progress	315
11.6 Comparisons of developments in Gāndhārī and other MIA languages	316
11.7 Scribal variations	317
11.8 Relationship to Dardic languages.....	318
11.9 Sociolinguistic considerations	320
11.10 Suggestions for further work	321
Appendix A1: Retroflex consonants in OIA	323
Appendix A2: <i>Anusvāra</i>	327
Appendix A3: Sanskrit consonant clusters	329
Appendix A4: Sanskrit syllable patterns	334
Appendix B: Perceptual cues of consonants	339
Appendix C: Comparative reflexes of Sanskrit consonants in Aśokan inscriptions	341
References	346

Maps and Tables

List of Maps

Map 1: Geographical location of the Greater Gandhāra language region (main map) and historical Gandhāra (enlarged inset map)	2
---	---

List of Tables

Table 3.4: Locations, Types, and Historical Eras of Gāndhārī source materials	37
Table 3.6: Summary and classification of source materials – third century BCE to third century CE	47
Table 4.1: The Sanskrit Phonological System	50
Table 4.4.1: Scale of relative sonority – from least at top to most sonorant at bottom	55
Table 5.0: The Gāndhārī Phonological system showing a representative version of the Kharoṣṭhī script	62
Table 8.1: Source document reference abbreviations (Abbr.) and texts	96
Table 8.2.1.2a: Gāndhārī reflexes of OIA intervocalic stops – Aśokan and BL sources	101
Table 8.2.1.2b: Gāndhārī reflexes of OIA intervocalic stops – BL, BC and RS sources	102
Table 8.2.1.2.6a: Common changes to single intervocalic unaspirated stops in Gāndhārī	125
Table 8.2.1.2.6b: Common changes to single intervocalic aspirated stops in Gāndhārī	126
Table 8.2.1.2.6.3: Suggested representation of Gāndhārī reflexes of OIA intervocalic stops in the lemma forms of words	133
Table 8.2.3.2a: Gāndhārī reflexes of OIA intervocalic fricatives – all sources	141
Table 8.2.3.2b: Overview of intervocalic fricative reflexes in Gāndhārī	145
Table 8.2.4.2a: Gāndhārī reflexes of OIA intervocalic semi-vowels – all sources ..	147
Table 8.2.4.2b: Overview of intervocalic semi-vowel reflexes in Gāndhārī	151
Table 8.2.6.2: The Gāndhārī Phonological system – suggested representation of word initial and medial single consonants	156
Table 8.3.1.1a: Gāndhārī reflexes of stop + stop clusters including geminates	159
Table 8.3.1.1b: Summary of stop + stop cluster developments in Gāndhārī	164
Table 8.3.1.2a: Gāndhārī reflexes of stop + nasal clusters	165

Table 8.3.1.2b: Summary of stop + nasal cluster developments in Gāndhārī	169
Table 8.3.1.3a: Gāndhārī reflexes of stop + fricative clusters	170
Table 8.3.1.3b: Summary of stop + fricative cluster developments in Gāndhārī	172
Table 8.3.1.4.1a: Gāndhārī reflexes of stop + <i>l</i> clusters	173
Table 8.3.1.4.1b: Summary of stop + <i>l</i> cluster developments in Gāndhārī	174
Table 8.3.1.4.2a: Gāndhārī reflexes of stop + <i>v</i> clusters	175
Table 8.3.1.4.2b: Summary of stop + <i>v</i> cluster developments in Gāndhārī	178
Table 8.3.1.4.3a: Gāndhārī reflexes of stop + <i>y</i> clusters	178
Table 8.3.1.4.3b: Summary of stop + <i>y</i> cluster developments in Gāndhārī	184
Table 8.3.1.4.4a: Gāndhārī reflexes of stop + <i>r</i> clusters	185
Table 8.3.1.4.4b: Summary of stop + <i>r</i> cluster developments in Gāndhārī	189
Table 8.3.2.1a: Gāndhārī reflexes of nasal + stop medial clusters	190
Table 8.3.2.1b: Summary of nasal + stop cluster developments in Gāndhārī	194
Table 8.3.2.2b: Summary of nasal + nasal cluster developments in Gāndhārī	195
Table 8.3.2.3a: Gāndhārī reflexes of nasal + fricative clusters	196
Table 8.3.2.3b: Summary of nasal + fricative cluster developments in Gāndhārī ...	197
Table 8.3.2.4.2a: Gāndhārī reflexes of nasal + <i>v</i> clusters	198
Table 8.3.2.4.2b: Summary of nasal + <i>v</i> cluster developments in Gāndhārī	199
Table 8.3.2.4.3a: Gāndhārī reflexes of nasal + <i>y</i> clusters	199
Table 8.3.2.4.3b: Summary of nasal + <i>y</i> cluster developments in Gāndhārī	201
Table 8.3.2.4.4a: Gāndhārī reflexes of nasal + <i>r</i> clusters	202
Table 8.3.2.4.4b: Summary of nasal + <i>r</i> cluster developments in Gāndhārī	203
Table 8.3.3.1a: Gāndhārī reflexes of fricative + stop clusters	204
Table 8.3.3.1b: Summary of fricative + stop cluster developments in Gāndhārī	213
Table 8.3.3.2a: Gāndhārī reflexes of fricative + nasal clusters	215
Table 8.3.3.2b: Summary of fricative + nasal cluster developments in Gāndhārī ...	219
Table 8.3.3.4.1a: Gāndhārī reflexes of fricative + <i>l</i> clusters	221
Table 8.3.3.4.2a: Gāndhārī reflexes of fricative + <i>v</i> clusters	221
Table 8.3.3.4.2b: Summary of fricative + <i>v</i> cluster developments in Gāndhārī	224
Table 8.3.3.4.3a: Gāndhārī reflexes of fricative + <i>y</i> clusters	225
Table 8.3.3.4.3b: Summary of fricative + <i>y</i> cluster developments in Gāndhārī	228

Table 8.3.3.4.4a: Gāndhārī reflexes of fricative + <i>r</i> clusters	229
Table 8.3.3.4.4b: Summary of fricative + <i>r</i> cluster developments in Gāndhārī	233
Table 8.3.4.1.1a: Gāndhārī reflexes of <i>l</i> + stop clusters	234
Table 8.3.4.1.4a: Gāndhārī reflexes of <i>l</i> + semi-vowel clusters	236
Table 8.3.4.1.5: Summary of <i>l</i> + C cluster developments in Gāndhārī	237
Table 8.3.4.2.4a: Gāndhārī reflexes of <i>v</i> + semi-vowel clusters	238
Table 8.3.4.2.5: Summary of <i>v</i> + C cluster developments in Gāndhārī	240
Table 8.3.4.3: Summary of <i>y</i> + R cluster developments in Gāndhārī	241
Table 8.3.4.4.1a: Gāndhārī reflexes of <i>r</i> + stop clusters	242
Table 8.3.4.4.2a: Gāndhārī reflexes of <i>r</i> + nasal clusters	246
Table 8.3.4.4.3a: Gāndhārī reflexes of <i>r</i> + fricative clusters	248
Table 8.3.4.4.4a: Gāndhārī reflexes of <i>r</i> + semi-vowel clusters	251
Table 8.3.4.4.5: Summary of <i>r</i> + C cluster developments in Gāndhārī	254
Table 8.3.5: Gāndhārī reflexes of OIA three consonant medial clusters	256
Table 8.3.6.1: Typical Gāndhārī reflexes of OIA word initial two consonant clusters	262
Table 8.3.6.2: Typical Gāndhārī reflexes of OIA word medial two consonant clusters	264
Table 9.1.1: Comparative development of Gāndhārī (G), Pali (P) and Śaurasenī (Ś) initial consonant clusters	281
Table 9.2.2.1: Typical MIA reflexes of OIA medial consonant clusters in the Aśokan Prakrits, expressed in terms of manner assimilation	284
Table 9.2.2.2: Comparative development of Gāndhārī (G) and MIA Śaurasenī (Ś) medial consonant clusters	286
Table 9.2.3: Comparative development of intervocalic single and cluster consonants in MIA	288
Table 9.3.3: Gemination of biconsonantal clusters C ₁ C ₂ in Sanskrit	293
Table 10.1: Gāndhārī consonant cluster reflexes involving epenthesis	301
Table A3.1: Sanskrit consonant clusters	330
Table A3.2: Three or more consonant clusters in Sanskrit	333
Table A4.1: Sanskrit two consonant clusters in terms of manner and relative sonority	334

Table B1: The main perceptual cues of consonants	339
Table C1: Comparative reflexes: Aśokan intervocalic single consonants	341
Table C2: Comparative reflexes: Aśokan medial consonant clusters	342

Symbols

The following sets of symbols are used in this work. Some of the common bracketing conventions of (), [], { } and < > are used in both Gāndhārī and Linguistic scholarship, to convey meanings related to each discipline. Since these conventions may overlap, their contextual uses are indicated in the following lists.

Gāndhārī: the following symbols have been used in the transcriptions of Gāndhārī words copied from source documents:

- [] an unclear or partially preserved akṣara whose reading is not certain
- (*) a lost or illegible akṣara that has been conjecturally restored on the basis of the context
- . a missing portion of a partially legible akṣara.
- * placed before a word - a presumptive equivalent without any direct parallel in Sanskrit or Pali texts.

Linguistic: the following notations are used in **linguistic formulae** which summarise Gāndhārī phonological developments:

- > becomes
- / in the environment of
- the position of the target phoneme relative to the environment. Hence A > B / X_ Y means that A becomes B in between X and Y
- C any consonant, including nasals
Place: l = labial, d = dental, r = retroflex, p = palatal, v = velar, cor = coronal
 - C_p any palatal consonant
 - C_l any labial consonant
 - C_{non-l} any non-labial consonant
 - C_{cor} any coronal consonant
- T Stop
- N Nasal
- S Fricative
- R Approximants and taps = Sanskrit semi-vowels
- V any vowel
- () round brackets = optional elements
- { } braces = alternative elements. Hence {a, b, c, etc.} = a or b or c etc.
- {x, (y)} round brackets inside braces = less common element
- # word boundary
- . syllable boundary

- + morpheme boundary
- Ø null element or deletion
- [p] an IPA phonetic symbol representing a sound
- / *p* / an OIA or Gāndhārī phoneme, represented using the IAST system
- ≡ has been represented by
This symbol is used to indicate a group of variant reflexes that have been used to represent a particular Gāndhārī phoneme in the source materials. This is typically expressed in the format

G phoneme ≡ G reflexes

These reflexes may have been used by different scribes in different manuscripts or inscriptions, at different times, but are sometimes used by the same scribe or group of scribes.

The meaning of the term reflex, as used in this thesis, is described in the ‘Explanation of Terms’ section (p. xxiii).

Abbreviations

AG-G ^L & AG-G ^S	Two Gāndhārī Manuscripts of the Songs of Lake Anavatapta (Anavatapta-gāthā) (ed. Salomon 2008)
Amg.	Ardha-Māgadhi
Apa.	Apabhraṃśa
Aś-Ins	Aśokan Inscriptions
av	avadānas
Av-G	Gandhāran Avadānas (ed. Lenz 2010)
BC	Bajaur collection of Kharoṣṭhī manuscripts
BCE	Before Common Era
BHS	Buddhist Hybrid Sanskrit
BL	British Library collection of Kharoṣṭhī manuscripts
c	Century. E.g., 3 rd c BCE = third century BCE
CE	Common Era
cl.	closed syllable
cpd.	compound
cpv.	comparative
Dhp-G ^K	Gāndhārī Dharmapada (ed. Brough 1962)
Dhp-G ^L & PY-G	A New Version of the Gāndhārī Dharmapada and a Collection of Previous-Birth Stories (ed. Lenz 2003)
EĀ-G	Three Gāndhārī Ekottarikāgama-Type Sūtras (ed. Allon 2001)
G	Gāndhārī
GBT	Gandhāran Buddhist Text series
GBM	Gandhāran Buddhist Manuscripts
GD	A Dictionary of Gāndhārī (Baums and Glass 2002)
Gos-G	An Edition and Study of the Goṣiga-sutra, the Cow-Horn Discourse (Silverlock 2015)
IAST	International Alphabet of Sanskrit Transliteration
IPA	International Phonetic Alphabet
Kh.	Kharoṣṭhī script
Khvs-G	A Gāndhārī Version of the Rhinoceros Sutra (ed. Salomon 2000)
Mah.	Mahārāṣṭrī
Mg.	Māgadhi
MIA	Middle Indo-Aryan
MS	Schøyen and other collections of Bamiyan Kharoṣṭhī manuscripts
MW	A Sanskrit English Dictionary (Monier-Williams 1899)
ND	Niya Documents
NIA	New Indo-Aryan
Nid-G	A Gāndhārī Commentary on Early Buddhist Verses (Baums 2009)
OIA	Old Indo-Aryan
P	Pāli
PIA	Proto Indo-Aryan

PIE	Proto Indo-European
PIIr.	Proto Indo-Iranian
Pkt.	Prakrit
PTSD	Pali Text Society Dictionary (Rhys Davids 1921-1925)
py	pūrvayogas
RS	Robert Senior collection of Kharoṣṭhī Manuscripts
SĀ-G	Four Gāndhārī Saṃyuktāgama Sūtras (ed. Glass 2007)
Śaur.	Śaurasenī
SC	Split collection of Kharoṣṭhī manuscripts
sd	scholastic documents
Skt.	Sanskrit
uno.	unless noted otherwise

Grammatical terms

Substantives

nom.	nominative
acc.	accusative
instr.	instrumental
dat.	dative
abl.	ablative
gen.	genitive
loc.	locative
voc.	vocative
dir.	direct (nom. and acc. cases together)
m.	masculine
f.	feminine
n.	neuter
sg.	singular
pl.	plural

In combination: e.g., m.nom.pl = masculine noun, nominative case, plural number

Verbs

pres.	present
caus.	causative
den.	denominative
impv.	imperative
1p. / 2p. / 3p	1 st person / 2 nd person / 3 rd person
abs.	absolutive

In combination: e.g., pres.3p.sg = present tense, 3rd person, singular

Other

adj.	adjective
adv.	adverb
ind.	indeclinable
pt.	participle
pp.	past participle
ppp.	past passive participle

Explanation of Terms

Old Indo Aryan (OIA): refers to languages of the Indo-Aryan family which existed in the regions of ancient India, Pakistan and Afghanistan between about 600 BCE and 1000 CE (Masica 1993: 50-51), known today through Vedic and Classical Sanskrit.

Vedic: Vedic Sanskrit, the Sanskrit of the Vedas

Classical Sanskrit: The Sanskrit prescribed by the Pāṇinian grammar.

Sanskrit (Skt.): refers to the collective of Vedic and Classical Sanskrit.

Although the terms OIA and Sanskrit are used interchangeably in the text at times, the use of OIA is typically intended to indicate the group of OIA languages in general, or general features of these languages, while Sanskrit is used when referring to particular OIA words or phonological features known through the Sanskrit record.

The **MIA languages** are often referred to as **MIA Prakrits**, or just **Prakrits**, and these terms are often used interchangeably. At a fundamental level the term Prakrit conveys the sense of a natural or spoken language, and has been used in India to convey a distinction with the more refined language of Sanskrit. At the same time there are many nuances in the use of this term. For a more detailed discussion refer to Ollett (2017; 11-14).

Reflex: the term reflex, as used in this thesis, typically refers to orthographic forms, recorded in the Kharoṣṭhī script for the Gāndhārī language, which are derivatives of earlier known OIA forms. While in many cases the phonological targets of these reflexes are clear, variant orthographic reflexes of single OIA forms are often recorded in the source materials. The phonological and phonetic targets of these variant reflexes are in some cases vague, multiple, or uncertain.

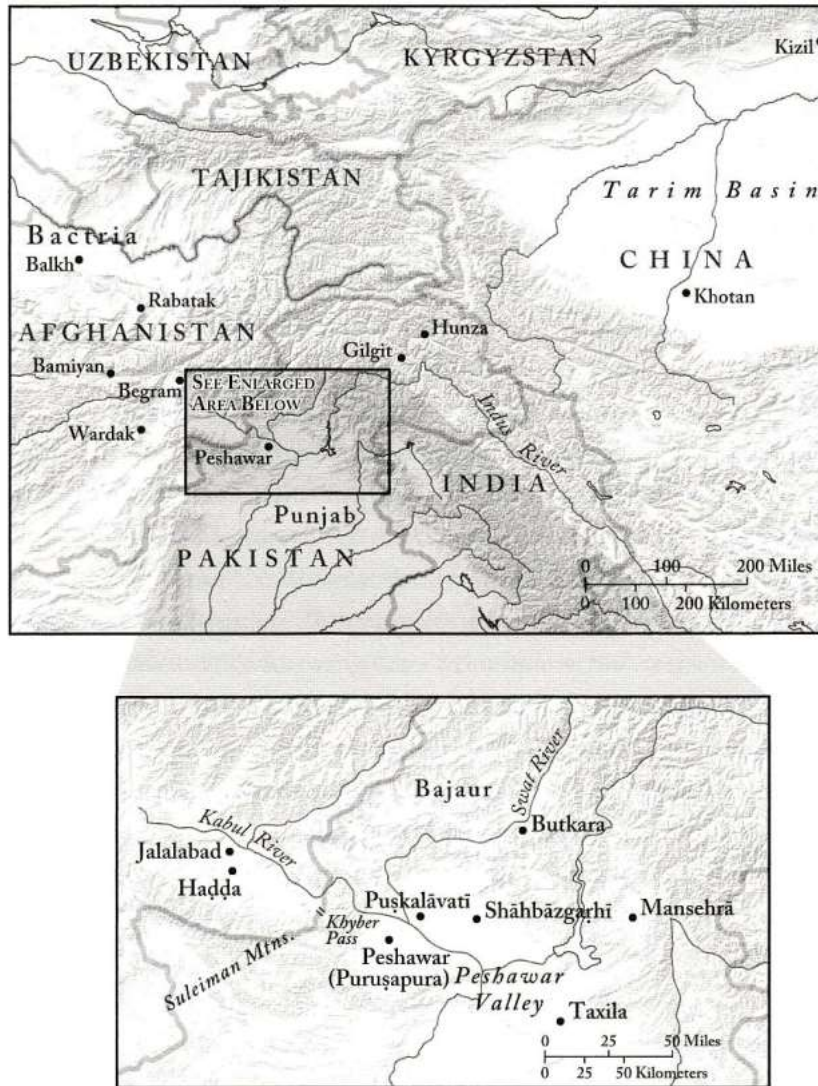
1. Introduction

1.1 Consonant phonology of the Gāndhārī language

This thesis documents and examines the consonant phonology of the Gāndhārī language. Gāndhārī is an ancient Middle Indo-Aryan (MIA) language which emerged as the dominant dialect in the Gandhāra region, encompassing areas of modern-day northern Pakistan and northeastern Afghanistan. Its use is also attested in a wider area including Bactria and southern Uzbekistan further to the northwest, as well as in the Central Asian Shan-Shan kingdom bordering the southern rim of the Tarim Basin in modern-day China. Gandhāra was the ancient name used for the Peshawar valley region, and the wider region of language use, including Bactria and the Central Asian areas of use, is also referred to as Greater Gandhāra in modern scholarship (see Map 1). The use of Gāndhārī is attested from the late fourth century BCE up to approximately the fourth century CE, through sources which include coin legends, inscriptions on rock, sculptures and reliquaries, administrative documents, and a growing number of recovered Buddhist manuscripts. These are typically written in the Kharoṣṭhī script, which was developed from the Aramaic script for use with the MIA Gāndhārī language in use in the area at that time¹.

In this thesis the consonant phonology of Gāndhārī, as recorded in source materials from the Gandhāra region, is collated and reviewed from a linguistic perspective, and several aspects of the phonology are examined in more detail. These include its relationship to the phonology of other MIA languages and the Old Indo-Aryan (OIA) languages as represented by Sanskrit; the variable recording of the phonology in the Kharoṣṭhī script, which provides a valuable record of sound and language change in progress; and the partially developed patterns of manner assimilation seen in Gāndhārī consonant clusters, compared with the more advanced nature of these developments seen in other MIA languages. Consideration is also given to some of the sociolinguistic factors which may have influenced these developments. A system for categorising the variable phonemic reflexes has also been proposed, with the aim of developing a practical shorthand method for recording these patterns.

¹ Salomon (1999: 2-5, 110-114; 2018: 11-12); Burrow 1937: v-vi



Map 1: Geographical location of the Greater Gandhāra language region (main map) and historical Gandhāra (enlarged inset map)

Source: Copied from Salomon (2018: Map 1, p. 10)

1.2 An outline of previous scholarship

Early knowledge of the Gāndhārī language and Kharoṣṭhī script came about through the decipherment of coins and inscriptions in the late nineteenth century. A wider knowledge of Gāndhārī became available through the translation and publication of other Kharoṣṭhī materials which became known or available to scholars around the same time. These include the Niya Administrative Documents, which are hand-written on wood and leather and were recovered in the Central Asian city of Niya by Aural Stein (Burrow 1937, 1940); the Khotan Dharmapada, which is written on birch bark scrolls and was recovered by Europeans in the Central Asian city of Khotan, although its origins are otherwise

unclear²; and the Aśokan rock inscriptions at Shāhbāzgarhī and Mānsehrā in northern Pakistan, which have been published within wider compilations of the Aśokan inscriptions³. Further early material included a compilation of non-Aśokan Kharoṣṭhī inscriptions inscribed on various base materials prepared by Konow (1929), which has subsequently been elaborated and expanded on by later scholars⁴. These inscriptions are typically short and repetitive with a limited vocabulary, but nonetheless provide a useable linguistic resource. Through these and other early works Gāndhārī became well known as a MIA language with many similarities to the other MIA Prakrits and Pali, and was originally known as the North Western Prakrit by late nineteenth and early twentieth century scholars.

Despite the recovery and translation of these materials, the extent of available Gāndhārī source materials remained limited until the early 1990s. Since that time, a significant number of Gāndhārī Buddhist texts, written on birch bark scrolls, have been discovered in the region of ancient Gandhāra and made available to scholars. In many cases these were found or acquired as collections of manuscripts which have become known by names associated with their current ownership, such as the British Library (BL), Robert Senior (RS), and Schøyen (MS) collections, some aspect of their provenance, such as the Split Collection (SC), or the location where they were found, such as the Bajaur Collection (BC). Comprehensive studies of a number of these manuscripts have since been published, including through the Gandhāran Buddhist Text (GBT) series⁵, or as PhD dissertations⁶, with further materials under preparation. While these publications have been prepared and presented in the wider context of Buddhist studies, similar to the earlier works of Hultzs, Konow, Burrow and Brough, they all contain detailed descriptions of the Gāndhārī grammar and phonological system as represented in the text of the published inscription or manuscript. In particular, the apparent changes which occurred between an assumed OIA ancestor language and the Gāndhārī Prakrit in which the documents were written are described, and extensive word indexes listing both Sanskrit and Pali equivalents, are also included in the more recent publications describing Buddhist

² Senart 1898, Brough 1962.

³ E.g., Cunningham 1877; Michelson 1909; Woolner 1924; Hultzs 1925.

⁴ E.g., Mehendale 1948, Fussman 1989, Falk and Karashima 2003; Falk and Karashima 2008; Baums 2012; McCrabb 2021.

⁵ Salomon 2000; Allon 2001; Lenz 2003; Glass 2007; Salomon 2008; Lenz 2010; Schlosser 2022; Cox (forthcoming).

⁶ Baums 2009; Lee 2009; Silverlock 2015; Schlosser 2016; Marino 2017; Butcher 2020.

manuscripts. Taken together these form an extremely valuable and diverse linguistic resource. A number of papers describing certain aspects of each collection, or features of some of the manuscripts within these collections, have also been published in various journals and conference publications⁷. While these provide valuable insights which add to the accumulating knowledge of Gāndhārī, by their nature they do not contain the same extent of phonological analysis and lexical listings as the GBT and PhD publications. Certain features of the Gāndhārī language and Kharoṣṭhī script have also been studied and published, both in thesis and journal formats⁸, while aspects of the phonology of the Gandhāran Aśokan inscriptions are presented in some recently published papers⁹. An outline of the Gāndhārī grammar, including a brief summary of the phonological developments reported in earlier scholarship, has been distributed on the Creative Commons by Baums (2019). Most of these developments are described in more detail in the earlier PhD dissertation by Baums (2009), and the observations contained in this publication have been referred to and discussed in the appropriate sections.

Publications in book format includes Richard Salomon's introduction to the British Library collection (Salomon 1999), which describes the discovery and early studies of this collection of manuscripts in the early 1990s, and a more recent publication which provides an accessible overview of the extent and nature of the Buddhist literature of ancient Gandhāra (Salomon 2018), as revealed through the growing body of available manuscript material. This volume also includes a good overview of the regional history and development of Buddhism in the Gandhāran region. At least two websites have been established which contain multiple links to Gāndhārī dictionaries, references, transliterations and digitized versions of inscriptions and manuscripts, and these offer an invaluable research resource¹⁰.

1.3 Aims of this thesis

The linguistic features of the Gāndhārī language contained in the materials published to date relate to individual inscriptions and manuscripts, which are in turn related to the work of individual scribes, locations, and eras. At the same time the

⁷ E.g., Salomon 2003a; Allon 2007; Allon et al. 2006; Strauch 2008a; Strauch 2008b; Falk 2011; Falk and Karashima 2012; Falk and Karashima 2013; Allon 2014; Falk and Strauch 2014; Salomon 2014; Strauch 2014; Falk 2015; Baums et al. 2016; Jatrasrisalai et al. 2016; Schlosser and Strauch 2016; Allon 2020.

⁸ E.g., Glass 2000; Strauch 2012b.

⁹ E.g., Suzuki 2019, 2020, 2021.

¹⁰ E.g., Baums and Glass 2002a; Allon and McCrabb 2019.

collective body of work represents a significant database of linguistic information about the Gāndhārī language, but one which does not yet appear to have been collated and studied as a whole within a linguistic framework. The primary aim of this thesis is to draw the material related to the consonant phonology together in a single volume, and provide a more detailed analysis and overview of the consonant phonology of the language. This work is developed around the following main areas of analysis:

- Collation of material from all the major published works into a systematically organised tabulated format.
- Analysis of the phonological and sound changes that occurred in the language compared with an assumed OIA ancestor language. One feature of the Gāndhārī phonology is the use of variable orthographic forms in the Kharoṣṭhī record to represent reflexes of particular OIA consonants or consonant clusters, which are assessed as part of this analysis.
- Development of a shorthand method for categorising and recording the attested forms and their variants, which facilitates a simpler representation of these variations and analysis of the associated sound changes.

The amalgamation and organisation of this material has allowed several other linguistic features of the language to be explored in more detail:

- The variable representations of consonant reflexes in the Kharoṣṭhī script are indicative of processes of sound and language change occurring in Gāndhārī at the time. This represents an invaluable record of language change in progress in a 2,000-year-old language.
- While the MIA languages all show similar patterns of development compared with the OIA ancestor languages, the extent and detail of these changes varies on a geographic and diachronic basis, and the Gāndhārī developments are evaluated in this context.
- A pattern of manner assimilation of the OIA consonant clusters occurred in the MIA languages. While fully developed in Pali for example, these changes are only partially developed in Gāndhārī, and an evaluation of these differences provides insights into the processes of manner assimilation.
- While the Gāndhārī morphology is not analysed, some of the effects of the phonological changes on the morphological structure are reviewed.

A comparison with the phonological system of Torwali, a modern day Dardic language of the Gandhāran area¹¹, has also been included to assist in understanding the presence, or otherwise, of any historical linguistic connection with Gāndhārī. Some of the potential sociolinguistic influences on the development of the MIA languages and Gāndhārī are also explored. These include the emergence of lingua franca languages, based on the dominant OIA dialects, as migrating Indo-Aryan speakers interacted with the indigenous language speakers of South Asia during the early millennia BCE; the simplifications of the OIA grammatical forms in these lingua francas; and the development of these lingua francas as regional vernaculars and early MIA Prakrits.

By amalgamating the previous Gāndhārī material in a single volume, and structuring this within a linguistic framework, it is intended to make this record more accessible and useable for Gāndhārī scholars, while updating or offering different approaches to some of the earlier linguistic interpretations. At the same time, it is intended that this work offers a Gāndhārī language resource which is accessible to the linguistics community. Naturally this project relies on and builds from the extensive work of others in this field, and nor is it intended to provide a definitive account of the Gāndhārī consonant phonology. The methodologies used in the examination of different aspects of the Gāndhārī phonology are described in more detail below.

1.4 Methodologies

1.4.1 Collation and Summary of materials

The phonological material presented in existing publications demonstrates differences between the OIA consonants and consonant clusters, and those in Gāndhārī, through the mapping of cognate word forms in Sanskrit, Pali and other MIA languages. Parallel MIA inscriptions exist for the Aśokan inscriptions for example, while Sanskrit and Pali cognates have been sourced from parallel texts in Buddhist manuscripts where available, or by interpretation in texts for which direct parallels are not available. These changes are typically categorised according to the OIA phonological structure, and word examples demonstrating these changes are either included in the text or sometimes listed separately. This material has been drawn together in a tabulated format in this thesis, organised on a consonant by consonant and cluster by cluster basis, categorised according

¹¹ Torwali is a Dardic language spoken in the Upper Swat valley of northern Pakistan.

to the manner and place of the consonants and consonant clusters. Tabulations identify the source materials, Buddhist scribes where known, and the approximate era of the sources. Typical word forms are listed for each pattern of change, recording the phonological development from an assumed OIA cognate to the Gāndhārī equivalent. It is intended that these tabulations present the material in a more accessible and consistent format, while allowing easier comparison between manuscripts, scribes and eras. More detailed analysis of various aspects of the language is also facilitated through these collated materials.

1.4.2 Analysis of phonological change between OIA and Gāndhārī

Language change develops around variations in sounds, and understanding the Gāndhārī phonological system relies on interpretation of the sounds represented by the Kharoṣṭhī script. As noted in § 1.1, this was developed from the Aramaic script for use with the MIA Gāndhārī language, and as such it does not accurately convey all phonetic features of the Indian phonology, for example in distinguishing between short and long vowels. This is further complicated by the recording of variant forms of reflexes for many of the OIA consonants. These may be recorded using either historic (conservative) or advanced (innovative) forms, which often appear as allographs in free variation, even in the same manuscript, or by the marking of *akṣaras*¹² with diacritics which suggest that scribes were attempting to record phonetic developments in the sounds. These may represent either the natural allophones of phonemes in particular environments, or possibly sounds that were transitional between conservative and progressed forms for which regular consonantal *akṣaras* were not available. It is important to determine what is being represented or intended by these orthographic variants, where a mixture of historic, transitional, and advanced spellings are recorded in many cases. The phonological and phonetic targets of these orthographic variants can in some cases be vague, multiple or uncertain, and interpretation often relies on contextual understanding. The variable reflexes are recorded in the transliterated versions of texts in the source materials, and care has been taken to transcribe all of these variations in the phonological record in the tabulated collations.

Single and clustered consonants are examined separately and grouped according to their manner and place of articulation, as well as word position. For each grouping an

¹² This is the commonly used Sanskrit term for graphemes and has been adopted in this thesis.

overview of the observed patterns of change is presented, followed by a section in which the linguistic factors associated with the changes are discussed, including comparison with cross linguistic typologies and a review of perceptual and phonetic factors associated with the change, while interpretations presented by previous authors are also reviewed.

Tabulations which summarise the overall patterns of change are included at the end of each section, and include simple phonological change rules, an assumed pronunciation for the Gāndhārī reflexes, and recommendations for an appropriate phonemic representation of reflexes in cases where alternate *akṣaras* have been recorded. A simple shorthand representation of the alternate reflexes is also provided, using the system described in § 1.4.3, offering an easy to read summary of the overall patterns of consonant reflexes found in the Kharoṣṭhī record. A linguistic explanation of some of the common types of sound change which occurred in Gāndhārī has also been provided.

1.4.3 Assessment and representation of variant Gāndhārī reflexes

A shorthand system for categorising and recording the variant forms of reflexes has been developed specifically for this thesis. The reflexes are assessed as to whether they most likely represent historic, transitional, or advanced spellings, and the forms are marked with shorthand superscripted tags to indicate these assessments. The categories assigned in this system are open to future interpretation, and the main aim is to present a system which is suitable to this purpose. An associated problem is selecting which of the variant reflexes, if any, represents the most likely language development, and ways in which a most appropriate form of representation for each the Gāndhārī phonemes might be selected from among the variety of recorded forms are explored.

1.4.4 The Gāndhārī record of language change in progress

The significant number of written source documents available within the Indo-Aryan (IA) literary and scriptural traditions, extending across nearly two millennia, has provided a rich resource for historical linguists seeking to understand the diachronic development of the IA languages¹³. Language change typically begins with the variation of sounds, as slightly different ways of saying the same thing enter a language (Campbell 2013: 193). While written material belonging to different eras provide a record of

¹³ The literature in this area is extensive. By way of example, Masica (1991) presents a comprehensive survey of the IA languages group, while Misra (1967) demonstrates the development of the historical phonology from the OIA languages through to modern standard Hindi.

diachronic developments in the phonology, documents belonging to a single era may also record both variation and change in different ways. Among the different types of historical document available to linguists for the study of language variation and change, such as direct and indirect transcripts, private letters, contemporary commentaries and literary sources, the latter are most limited in providing a record of the variations that arise in real-life speech events (Schneider 2002: 70-81). At the same time, they can still provide insight into phonological features that are undergoing change, or disappearing in a dialect, which are sometimes “systematically represented by non-standard dialect orthography as written by native speakers” (Schneider 2002: 80). It has also been suggested that variable representations tend to be functional to some extent, and that non-standard orthography is often phonetically based (Schneider 2002: 88). Similar to this ‘literary’ category, the administrative, inscriptional and scriptural documents which make up the Gāndhārī source materials are formal documents which may not fully reflect the natural speech of the scribes, who most likely spoke different dialects or even mother tongues, while the scriptural language also incorporated various eastern MIA influences along with apparent Sanskritisation from around the third century CE. On the other hand, Gāndhārī originated as a local dialect which developed into a *lingua franca* of the diverse population of the Gandhāra region in the late centuries BCE, coming to serve as the scriptural language of the Gandhāran Buddhist communities in the early centuries CE (Salomon 2018: 62). In this sense the language used in these source materials is closely related to the living vernacular language in its home region, and was undoubtedly close to the language spoken by many of the scribes preparing these materials.

The most extensive body of Gāndhārī source materials is provided by the first and second century CE Gandhāran Buddhist manuscripts (§ 3), and aside from comparisons with the third century BCE Asokan inscriptions and some later stage third century CE manuscripts, these effectively provide a synchronic record of the language. Gāndhārī was almost universally recorded using the Kharoṣṭhī script¹⁴, and within this body of Kharoṣṭhī manuscripts there is a significant variability in the representation of phonemes. This includes the use of both historic and advanced forms as allographs used in free variation, along with diacritic modified *akṣaras* which appear to represent allophones or transitional

¹⁴ Kharoṣṭhī and Brāhmī are the earliest scripts that were used widely in ancient India, and are first reliably attested in the mid third century BCE Aśokan rock inscriptions. Both were developed for the recording of MIA rather than OIA languages (Salomon 2003a).

pronunciations (e.g., Strauch 2012). This suggests that along with the diachronic record available through comparison of the third century BCE and first and second century CE materials, the later literary materials offer a valuable synchronic record of the changes which were occurring in the language at the time. Such variable orthography may have developed for several reasons. It has been observed that word alternations often occur in early scripts having an undeveloped orthography (e.g., Lahiri 2015), and it is quite possible that the Gāndhārī scribes had not developed a clear understanding of the differences between a phoneme and its allophones, for example (Fussman 1989: 467 [§ 27.1]). This would undoubtedly have been compounded by the cross-cultural nature of the script, which had not conventionalised all aspects of the Indian sound system¹⁵. The recording of many advanced forms also suggests that these had been developing in the language for a considerable time prior to the introduction of the written record, while the orthography may also reflect different levels of dialectic development between communities of the Gandhāra region¹⁶. Phonetic remnants of language variation can remain embedded in a spoken language, and manuscripts may incorporate many of the changes which occurred in the language's sound system prior to its script coming into existence¹⁷. They may also retain historic spellings of earlier pronunciations used prior to standardisation of spelling systems, seen in the English vowel shift cases of *ceme* / *came* and *symed* / *semed* 'seemed' for example (Campbell 2013: 398). The use of both historic and advanced allographs in Gāndhārī may also reflect some of the trans-Indian MIA affiliations of the different Buddhist communities and schools. Working as they were within this complex linguistic and orthographic situation, and in such a diverse cultural region, the Kharoṣṭhī scribes appear to have worked with a degree of hesitancy or uncertainty about the actual phonological structure of the language and its best representation. Consequently, the language was recorded as much at a phonetic level in relation to the vernacular language, as through any precise phonological system, while various allographs which reflect diachronic, dialectic or trans-Indian variations appear to have been used inconsistently.

¹⁵ The Kharoṣṭhī script is by its nature more phonetic and less phonological than Brāhmī, which was used more widely across northern India (Strauch 2012: 148; Salomon 2018: 62).

¹⁶ The variations which arise between dialects used in central or high contact areas of language use, and those in remote or insular areas, are well documented (Trudgill 2011: 3-9).

¹⁷ Gogoi, Morey and Pittayaporn 2020. This is an interesting case study of allographic patterns seen in 17th and 18th century texts in the Tai Ahom language, recorded on sasi bark manuscripts. More information about the project to document, conserve and archive these manuscripts can be obtained from the following website: <https://eap.bl.uk/project/EAP373>

While not all variability and heterogeneity in a language structure is associated with change (Weinreich, Labov and Herzog 1968: 187-188), the extensive nature of these features in the Kharoṣṭhī record suggests that they provide insights into patterns of change in progress in the Gāndhārī language. As expressed by Matsumoto (2019: 199):

“A change from an archaic form A to an innovating form B, regardless of whether they are continuous or discrete, always involves an intermediate stage where A and B coexist in the speech of individuals, in the speech of communities and in the structure of language”.

In this context, the two thousand year old Gāndhārī inscriptional and manuscript materials offer a valuable record of the processes of language variation and change that had been, or were occurring, in the ancient language and dialects of the Gandhāran region. The method described in the previous section for the classification and recording of the alternate reflexes provides a simple system for representing these patterns of change, which have been collated and tabulated in this thesis for ease of reference.

1.4.5 Comparative development with other MIA languages

The differences that arose between the various regional MIA languages, including Pali, as they developed from the OIA ancestor languages, are well documented¹⁸, although comparative knowledge of Gāndhārī was limited by the relative scarcity of available source materials prior to the early 1990s. The record of Gāndhārī source materials has increased considerably since then through the body of newly discovered Buddhist manuscripts. These provide a wider and more substantial reference point for the comparison of Gāndhārī with other MIA Prakrits, throwing further light on its place in the comparative development of the MIA languages. These comparisons are reviewed in relation to the consonant developments analysed in this thesis, and in particular, the comparative diachronic development of the single intervocalic consonants, consonant clusters and sibilants. While some developments appear to be more progressed in Gāndhārī than in other branches of MIA, others are relatively less progressed, raising interesting questions about the comparative influences which may have affected these patterns. Some of these influences are undoubtedly sociolinguistic, and while it is

¹⁸ E.g., Pischel 1879; Woolner 1917; Geiger 1943; Misra 1967; Salomon 1998b; Bubenik 2003; Oberlies 2019.

impossible to ascertain the detailed nature of these influences on the development of the MIA languages, some of the possible factors are reviewed.

1.4.6 Manner assimilation of OIA consonant clusters in MIA

One of the interesting linguistic features seen in the MIA languages is the consistent development of OIA consonant clusters as geminates of the stronger consonant, through both regressive and progressive manner assimilation. While this assimilation is fully developed in all of the Pali clusters, as well as the later literary Prakrits, it is only partially developed in the Gāndhārī record. Unlike place assimilation, which is widespread in language change and mostly develops in a consistent manner on a cross linguistic basis, manner assimilation appears to be recorded less often, and has also been observed to spread less reliably¹⁹. Linguistic and phonetic explanations for MIA manner assimilation are reviewed, and the comparative levels of development between the Gāndhārī and Pali assimilation explored in this context, providing insights into the development and variabilities of manner assimilation. The partial development of the changes to clusters in Gāndhārī also provides a valuable snapshot of the staged progression of these assimilatory changes in the MIA languages.

1.5 Overview of chapters

The aim of this thesis is to explore certain aspects of the Gāndhārī language, and present the material in a way which is of potential interest to both linguists and Gāndhārī scholars and students alike. In so doing it has been felt necessary to develop the thesis in a logical sequence that explores the material within an overall context, and content which explains some of the more foundational knowledge in each of these disciplines has been included, with the intent of providing background material for readers in the other. In some cases where the material is quite detailed this has been placed in Appendices. Readers in either discipline may find some sections unnecessary or superfluous, while others are more pertinent to their interest. An overview of the content of each chapter is provided here for guidance:

- Chapter 2 presents an overview of the Gāndhārī language including brief descriptions of the linguistic and sociolinguistic environment in which it developed and was used, and is intended to provide a general background for anyone unfamiliar with this area.

¹⁹ E.g., Padgett 1994: 465; Gordon 2016: 128-132.

- Chapter 3 describes the main source materials by which the Gāndhārī language is known, along with details of the publications from which the material has been analysed and studied. It has been necessary to focus on particular sources in this thesis, and the reasons for these selections are described.
- Chapter 4 presents an overview of the phonetics and phonology of Sanskrit. The OIA languages are known today through Sanskrit, which is the closest known language to the OIA ancestor language of Gāndhārī, and as such provides a basis from which to compare the phonological changes which occurred in Gāndhārī. Aspects which are of interest in understanding the development of MIA languages such as Gāndhārī are described, including the phonological system, the sandhi system, the spread of retroflexion, nasal consonants, consonant clusters, syllable structure and geminate consonants. Some of the more detailed material has been placed in the Appendices.
- Chapter 5 provides an overview of the Kharoṣṭhī script and Gāndhārī phonological system. While the script is cursive and shows a natural variability in the hands of different scribes, a set of typical grapheme characters which have been used to represent the consonants and consonant clusters are tabulated for reference, along with some of the more common diacritic and other forms of syllabic modifiers.
- Chapter 6 reviews some of the more common types of sound and language change which occurred in Gāndhārī and other MIA languages, drawing on explanations of the articulatory basis for some of these changes, and comparisons with patterns observed on a cross linguistic basis. These include place and manner assimilation within consonant clusters, and some of the more common patterns of change to single consonants. Some of the more detailed material has been placed in the Appendices.
- Chapter 7 describes some of the methods utilised in recovering the phonology of the Gāndhārī language. In addition to comparison with parallel materials in the OIA and MIA languages, interpretation of the orthographic intent of the Gāndhārī scribes is an important aspect in this work. Approaches to the phonemic representation of Gāndhārī phonemes in ways which incorporate the variable representations found in the source materials are discussed.
- Chapter 8 presents the collation, analysis, and review of the available material related to the changes in the OIA consonants and consonant clusters in Gāndhārī, for

consonants occurring in word initial, medial and final positions. This chapter forms the major component of the thesis.

- Chapter 9 examines the comparative development between Gāndhārī, Pali and other MIA languages. Differences between the development of intervocalic consonants, consonant clusters and sibilants are compared on both a synchronic and diachronic basis. A number of analytical approaches to understanding the strength based manner assimilation of consonant clusters are explored.
- Chapter 10 describes some of the common patterns of sound change observed in the Gāndhārī record.
- Chapter 11 presents the conclusions which may be drawn from the work.

2. The Gāndhārī language

2.1 Overview

Gāndhārī is an ancient Middle Indo-Aryan (MIA) language which emerged as the dominant dialect in the Gandhāra region, encompassing areas of modern-day northern Pakistan and northeastern Afghanistan. Its use is attested in a wider area including Bactria and southern Uzbekistan further to the northwest, as well as in Central Asian Shan-Shan kingdom bordering the southern rim of the Tarim Basin in modern-day China. Gandhāra was the ancient name used for the Peshawar valley region, and the wider region of language use is also referred to as Greater Gandhāra by modern scholars²⁰, while Gāndhārī was originally known as the North Western Prakrit by late nineteenth and early twentieth century scholars. Gāndhārī is attested today through source materials which include inscriptions on rock, sculptures and reliquaries, coin legends, administrative documents, and a growing number of Buddhist manuscripts, which date from the late fourth century BCE to around the third or fourth century CE. Due to its geographic location the Gandhāra region has been an important military, trade, and cultural crossroads throughout history, and during this period was ruled by a series of foreign invaders including Persians, Mauryans, Indo-Greeks, Scythians, Parthians and Kuṣāṇas. Many of these rulers became state sponsors of Buddhism, which emerged as an important unifying factor in the regional culture during this period, and Gāndhārī became a lingua franca of this culturally diverse population²¹. It was also carried into in the Central Asian Shan-Shan kingdom centering on Niya, Endere and the Lou-Lan area in the Tarim Basin, where it was used as an administrative language²².

Like other MIA languages, Gāndhārī developed through a series of changes to one of the Old Indo-Aryan (OIA) dialects, generally in the form of simplifications to the phonology and morphology (Masica 1991: 167, 231, 257), although it is unclear whether the spoken language may have also gained complexity in other areas of its use. The language in use up to about the first century CE appears to represent a relatively intermediate stage of MIA development. For example, most of the OIA intervocalic

²⁰ Refer to Map 1 in Section 1.1

²¹ Good accounts of the general historical and linguistic background of Gandhāra and the Gāndhārī language are given by Salomon (1999: 2-5) and Salomon (2018: 11-49, 62-68). These form the basis for much of what is presented in this section.

²² Burrow 1937: v-vi.

consonants have been weakened, compared with Pali where these tend to be preserved, and later stage dialects such as Mahārāṣṭrī where they have mostly been elided (Salomon 1999: 124-125). Such elisions also appear on a regular basis in some of the late first century CE Gāndhārī inscriptions (Fussman 1989: 457 [§ 18.3]), suggesting similar ongoing changes to the language. While such observation may point towards elements of common progression in the MIA languages, distinct regional features are also observed. These include retention of the three OIA sibilants in Gāndhārī, which are typically reduced to one sibilant in other MIA languages, and the retention of some consonant clusters, such as combinations of stops and semivowels for example, which are otherwise assimilated (Salomon 2002: 119-120).

All of the Gāndhārī source materials are written using the Kharoṣṭhī script, which appears to have been based on or adapted from the Aramaic script used by the Persian Achaemenid Empire which ruled Gandhāra from 559-336 BCE²³. It was possibly introduced as a device of cross-cultural convenience in order to represent the Indian sound system (Konow 1929: xiii), but at the same time not all aspects of this system were conventionalised within the script, such as the non-marking of Indian geminate consonants and vowel length (Strauch 2012: 136). This makes clear understanding of the Gāndhārī phonological system more difficult, and is compounded by the variable and at times seemingly confused representation of phones and phonemes by some scribes²⁴. While the origins of the Kharoṣṭhī orthography remain unclear, it most likely evolved through a multi-stage process rather than as a single invention, with progressive modifications of the Aramaic signs to represent different Indic phonemes. While the similarities between some graphemes such as Aramaic *bēt* (𐤁) and Kharoṣṭhī *ba* (𑀕) are reasonably clear, they are more obscure in others, as in *kap* (𑀭) and *ka* (𑀓) for example (Glass 2000: 11-20). The script was used on the legends of coins issued by most of the later foreign rulers of Gandhāra, including the Indo-Greeks and early Kuṣāṇas (Salomon 2018: 24-25, 42-43), indicating the continued use of Gāndhārī under these empires. The earliest significant record of the script is seen in the Shāhbāzgarhī and Mānsehrā inscriptions of the Mauryan emperor Aśoka located in northern Pakistan, dating from around 250 BC²⁵. While the

²³ Glass 2000: 1-2; Salomon 2018: 65.

²⁴ Strauch 2012: 148; Salomon 2018: 64-65.

²⁵ Glass 2000:1-2; Salomon 2018: 19-20.

origins of the Indian Brāhmī script are similarly uncertain, the earliest dateable records are also seen in the wider body of pan-Indian Aśokan rock and pillar inscriptions²⁶.

It is generally thought that Buddhism was introduced to Gandhāra during the reign of Aśoka around the middle of the third century BCE (Salomon 2018: 18). Buddhism was able to prosper under the royal patronage of later foreign rulers, and the production of written Buddhist manuscripts in the Gāndhārī language flourished in the first and second centuries CE during the reign of the Kuṣāṇa kings (Salomon 2018: 44-45). A growing collection of Gāndhārī Buddhist manuscripts from this period have been discovered during the past twenty or thirty years, which undoubtedly represents only a small part of a more extensive literary tradition²⁷. Many of these manuscripts were found in reliquary pots inscribed with details of the holy gift or relic. During this period Gandhāra became one of the major Buddhist centres in the Indian subcontinent (Salomon 1999: 5), and it is possible that the status of Gāndhārī was similar to that of Pali and Sanskrit among the Buddhist languages of the time (Salomon 2014: 1). Gandhāra sat at the centre of major trade routes from India to the west and north, and Buddhism was able to travel from there into Central Asia and further on to China (Salomon 2018: 48-49). Gāndhārī manuscripts and fragments have been found near old Silk Road cities in Xinjiang, including the so called Khotan Dharmapada which was reportedly discovered near Khotan on the southern route around the Tarim Basin, and some small manuscript fragments found by Paul Pelliot near Kucha on the northern route²⁸. Along with Kharoṣṭhī inscriptions found near early centres of Chinese Buddhism at Lo-yang and Chang-an, these suggest that Gāndhārī played an important role in the early transmission of Buddhism to these regions²⁹. An associated development was the use of Gāndhārī as an official language in the Shan-Shan or Kroraina kingdom during the third century CE, seen in administrative documents found close to Niya on the southern rim of the Tarim Basin by Aurel Stein, and more recently by Chinese archaeologists³⁰. Although very similar to the Gāndhārī used in Buddhist documents, the language of these materials contains many Iranian and other loan words. All of these developments suggest that Gāndhārī had become an important regional language during

²⁶ The origins of the Brāhmī script are similarly uncertain, and one suggestion is that it could have been loosely modelled on the Kharoṣṭhī script during the Aśokan period. For a discussion of the various theories on the origins of the Brāhmī script see Salomon 2003b: 85-89 and 2018: 65-66.

²⁷ Salomon (2018: 59). For an overview of the different genres of Buddhist literature contained in the recent finds of Gandhāran Buddhist manuscripts refer to Falk and Strauch (2014: 54-57).

²⁸ Brough 1969: 2; Salomon 1998a.

²⁹ Salomon 1999: 5-7, 58-59; 2018: 47-49.

³⁰ Burrow 1937; Salomon 1988; 1999: 113.

the early centuries CE, and at the same time appears to have preceded Sanskrit as a carrier of Buddhism into Central Asia (Brough 1962: 49).

Although spoken Gāndhārī undoubtedly continued to evolve, the tradition of Buddhist Gāndhārī literature, and with it knowledge of the evolving language, faded from view with the decline of Gandhāran Buddhism from around the fourth century CE³¹. While major transformations occurred between the MIA and New Indo-Aryan (NIA) languages, many details of this process remain unclear (Masica 1991: 55), and this is especially the case for Gāndhārī. Sufficient similarities have been observed to align Gāndhārī with the modern Dardic language group, however, and it has been most closely linked to Torwālī in the upper Swat valley³².

There is no known record of a grammar or prestigious text that served as a grammatical reference point for Gāndhārī (Fussman 1989: 486 [§ 37.4]), which raises questions about the way in which the language was understood at the time it became written. The variable spelling and phonemic forms used in Gāndhārī texts suggests on the one hand that scribes may have been attempting to understand and record the evolving sounds of the living language, unguided by any firmly established grammatical conventions. This compares with ancient scholars working with languages like Pali and Sanskrit, which had been standardized over time through careful grammatical analysis (Salomon 2018: 63-64). On the other hand, many of the Buddhist monks were undoubtedly familiar with Pali and other MIA languages, possessing sufficient knowledge to enable translation from the source texts in those languages. A familiarity with Sanskrit is evidenced in some of the later third century manuscripts, for example, where adjustments had been made in some cases to restore conservative Sanskritic features in an apparent process of Sanskritisation of the language (Salomon 2001). At the same time significant variability occurs in the grammatical forms recorded in Gāndhārī manuscripts, as for example in the variety of forms used for declensional endings, pointing towards a lack of any standardized approach (Salomon 2018: 64). It is also likely that monks would have come from different linguistic backgrounds, and the levels of knowledge were no doubt variable. It is possible that some monks may have been fully literate in several languages and able to translate texts, while others may have been literate in Gāndhārī only,

³¹ Salomon 2014: 1; 2018: 45-47.

³² Burrow 1937: vi; Fussman 1989: 439 [§ 3.1]; Masica 1991: 21.

able to work as copyists but possibly more prone to introduce vernacular forms into transcribed texts, while others may have been only weakly literate but practiced in the chanting of texts only. In the absence of records however, the extent to which the language was understood and organised grammatically, how well such knowledge was disseminated among the scribes, how this influenced the available manuscripts and the diachronic development of such knowledge, all remain unclear.

2.2 Linguistic environment

2.2.1 Historical background

The Indo-Aryan (IA) languages, which are spoken widely today throughout South Asia, are classified as a sub-branch of the Indo-Iranian branch of the Indo-European language family. The predominant view among Western scholars, based on carefully reconstructed linguistic family trees, and supported by some archaeological evidence, is that the original speakers of the ancestor Proto Indo-European (PIE) language lived somewhere in the region of the Black, Caspian and Aral Seas some 6,000 to 9,000 years ago. Among the many subsequent migrations and language developments which occurred during the spread of Indo-European speakers across Europe, the Middle East, and South Asia, it is believed that Indo-Aryan speakers migrated into the northwest of South Asia at some time around 1500 BCE, introducing their languages and culture to the region³³. It has been suggested that Proto Indo-Aryan (PIA) speakers had settled and become well established somewhere around the Iranian plateau prior to this time, and had divided into two groups, one to the West and the other in the eastern area closer to the border between Afghanistan and Pakistan³⁴. This latter group eventually began to migrate into the fertile Indus valley regions, potentially motivated by population pressures in their homelands. These migrations most likely occurred in several waves over a number of generations, and the Indo-Aryans migrating in subsequent waves appear to have spoken slightly different dialects³⁵. It is also possible that the PIA speakers had established both political and trade connection in the Indus valley regions prior to these migrations, and that a dialect of PIA could potentially have been in use in the northwest as a lingua franca from an earlier time. The number of Dravidian loanwords found in the earliest records of the Indo-Aryan

³³ E.g., Burrow 1972: 30-34; 1973; Masica 1991: 32-37; Clackson 2007: 1-19; Fortson 2010: 8-13, 39-49. The exact location and dates of the homeland are the subject of much scholarship and debate.

³⁴ Burrow 1973: 125-126.

³⁵ Burrow 1972: 32; 1973: 125-126; Norman 1992b: 225-228.

languages suggests that the Indo-Aryans possibly met and intermixed with the indigenous peoples of the Indus valley region from an early stage in their migrations (Norman 1992: 226-227), although the lexical evidence in this regard is not necessarily conclusive. The Khyber Pass, located in present day Pakistan, which has been used by invaders throughout recorded history, was most likely the main route of migration. This route extends the natural line of movement from the Aral Sea steppes, via the Oxus and Kabul River valleys, into the fertile plain of Peshawar. Other potential migration routes include the Bolan Pass, located in present day Pakistan, and other more southerly passes, bringing together different dialects of the migrating peoples and the indigenous people of the northwest.

2.2.2 The OIA languages

The OIA languages are the earliest attested IA languages, dating back to around the time of the Indo-Aryan migrations into South Asia, and known today through the oral and written forms of Vedic and Classical Sanskrit. Composition of the Vedic hymns appears to have commenced sometime around 1500 BCE, beginning with the Ṛgveda, and Vedic Sanskrit has been exceptionally well preserved through a continuous tradition of oral recitation through to the present day³⁶. The nature of the material and strength of the system of oral transmission within the Brahmin community meant the hymns were only recorded in written form at a much later time³⁷. The language of the Ṛgveda appears to have been an amalgamation of several earlier dialects, and is essentially an archaic or conservative form of OIA, potentially different to the spoken languages of the time³⁸. At the same time, a number of MIA or ‘Prakritic’ features can be discerned from the odd popular forms seen in the Vedic texts, suggesting that the development of these forms had begun before the Indo-Aryan migrations into South Asia³⁹, while a number of later outside influences, such as Dravidian and Munda loanwords, are also found in the language (Witzel 1999).

Due to the strong influence of the developing vernacular forms on Vedic Sanskrit, the Sanskrit grammarians progressively formalised the grammar of what became known as Classical Sanskrit, which culminated in the development of Pāṇini’s well-known

³⁶ Burrow 1972: 35-36; Witzel 1999: 2-5.

³⁷ The earliest writing systems used in India date from around the 3rd century BCE, and were developed for the recording of MIA rather than OIA languages. See Salomon 1995 and 2003b.

³⁸ Witzel 1989: 97; Oberlies 2019: 18-19.

³⁹ Norman 1992b: 228; Oberlies 2018: 18-19.

grammar, the *Aṣṭādhyāyī*, around the fourth century BCE. Classical Sanskrit did not derive directly from Vedic, but effectively preserved the language in use by the educated Brahmins of the northwest at that time⁴⁰. This represents a parallel but much slower change of the language within a conservative educated community, compared with the MIA forms spoken by the general population. While the two languages are very close, the development of Sanskrit literature between the time of the Ṛgveda and Classical Sanskrit reveals many areas where the grammar underwent change⁴¹. The grammars of both Vedic and Classical Sanskrit are well known in the Indian traditions, and descriptions of these grammars have been compiled and published by many Western scholars⁴².

2.2.3 The MIA languages and Pali

As noted above, development of the MIA languages can be traced from occasional innovative forms found in the Ṛgveda, including phonological, morphological, and syntactic variations, which are similar to those found in later recorded versions of MIA. These suggest the presence of one or more early ‘Vedic Prakrits’ which developed alongside Ṛgvedic Sanskrit. These already contained many features of the later MIA languages, which appear to be descendants of these varieties rather than simple continuants from the known OIA languages⁴³. The earliest significant records of the MIA languages are found in the regional dialects of the Aśokan inscriptions, Pali, the language of the early texts of Theravāda Buddhism, and Ardha-Māgadhī, the canonical language of the Jains, all dating from around the third century BCE⁴⁴. The Aśokan inscriptions, dating from around 250 BCE, are the first accurately dateable records of this linguistic development, and due to their geographic spread provide a reasonable view of the diversity of the MIA dialects at that time. Three main dialects are attested, which are best represented by the Shāhbāzgarhī and Mānsehra edicts in the Northwest, which are recorded in the Kharoṣṭhī script, the Gīrnār edict in Gujarat in the West and the Dhāuli and Jaugūḍa edicts in Orissa in the East, recorded in the Brāhmī script, while the so-called

⁴⁰ Burrow 1972: 36; Norman 1992b: 229-231; Oberlies 2019: 9-10, 35. Salomon (1999:4) observes that “the Sanskrit spoken in Gandhāra was held to be the purest and most refined form of the sacred language.”

⁴¹ E.g., Macdonell (1927: 236-244), Burrow (1972: 35-44) and Cardona 2003 (105-108) provide useful descriptions of these differences.

⁴² The list of such works is extensive. It includes early works such as those by Macdonell (1910) and Whitney (1924), along with more contemporary works such as those by Burrow (1972) and Oberlies (2003a) for example, among many others.

⁴³ Norman 1992b: 228-230, Oberlies 2019: 21-33.

⁴⁴ Masica 1991: 51-52; Oberlies 2019: 9-11, 22.

eastern forms tend to predominate in the central and southern inscriptions⁴⁵. Pali and Ardha-Māgadhī present as early manifestations of a common language of the Gangetic plain, which had developed from the Vedic Prakrit of the eastern region (Oberlies 2019: 9). Several stages of development are seen in the Pali canon, which has been dated from as early as the third century BCE, when tradition holds that the oral canon was brought to Sri Lanka⁴⁶. Pali appears to be a compilation of mixed dialects, with literary Pali most similar to the Aśokan dialect of Gīrnār in the west⁴⁷.

Although the MIA languages are often described as regional vernaculars or Prakrits, they are largely known through surviving inscriptional records and literary sources, which represent a narrower tradition of literary Prakrits (Ollett 2017: 11-16). In addition to the Aśokan inscriptions, other early inscriptional sources include a variety of non-Aśokan Brāhmī inscriptions of varying length and importance, dating from around the third century BCE through to the sixth century CE. These show variations on a regional basis, including an eastern group based upon inscriptions generally located in Uttaranchal and Bihar on the Gangetic plain, a central group from locations in Madhya Pradesh and Chhattisgarh in central India, a southern group from locations in Andhra Pradesh further to the south, and a western group from locations in Gujarat and Maharashtra in the west⁴⁸. Among the major literary Prakrits, Māgadhī is considered to have developed from an eastern dialect, Śaurasenī from a central dialect, and Mahārāṣṭrī from a south western dialect, while Śaurasenī appears to be the earliest of the literary Prakrits⁴⁹.

The dates of the Buddha's life are uncertain and controversial, although it is thought that he lived sometime between the sixth and fourth centuries BCE based on divergent records of his birth date recorded in the Sri Lankan and Indian Buddhist traditions. More recent scholarship has suggested his death occurred somewhere around 400 BCE⁵⁰, indicating that he taught during the fifth century BCE, about 150 years prior to the time of Aśoka. The Theravādin Buddhist tradition holds that the language spoken by the Buddha was Pali, the language of its written texts, which was also the same as

⁴⁵ E.g., Salomon 1998b: 133-140; Norman 1994; Oberlies 2003b: 163-167; Oberlies 2019: 42-43.

⁴⁶ Geiger 1943: 2-3; Katre 1964: 11-12; Oberlies 2019: 5. The early canon would undoubtedly have existed prior to this date. It was not committed to writing until a later stage around the late first century CE (§ 2.2.5)

⁴⁷ Geiger 1943: 2-3; Oberlies 2003b: 166-167.

⁴⁸ E.g., Mehendale 1948: xv-xvii; Salomon 1995: 140-146.

⁴⁹ Woolner, 1917: 5-6; Misra 1967: 103; Masica 1991: 51-53. There do not appear to be any clear records of the earliest literary Prakrits, but these appear to have emerged from just prior to or around the middle of the 1st millennium CE (E.g., Woolner 1917: 71; Masica 1991: 56; Bubenik 2003: 206).

⁵⁰ E.g., Cousins 1996; Salomon 2018: 16.

Māgadhī, an eastern language⁵¹. While it is likely that the Buddha spoke an eastern dialect, linguistically Pali presents as a western language however, seen through many important features shared with the western Aśokan edicts, especially at Gīrnār, but without the retained *Cr* clusters for example. And while it also carries some sporadic eastern features, other major features of the eastern Aśokan Prakrits are absent from Pali, as are most of the features of the northwestern dialect. One of the dialects from which Pali is derived appears to have had some affinities with the ancient Avestan language of Persia, and as a western language the ancestor dialect was possibly located near the borders with ancient eastern Iran, around the Bolan Pass area above modern-day Quetta in western Pakistan (Oberlies 2019: 21-22, 42-52). Like other MIA languages, Pali is not a direct continuation of the language of the Ṛgveda, but rather descends from one of the other OIA Vedic dialects. It is older than Classical Sanskrit and therefore does not derive directly from it (Oberlies 2019: 21, 35). The eastern features found in Pali appear to be traces of an eastern dialect from which the earliest Buddhist texts were possibly transcribed, and Pali possibly originated as a kind of *lingua franca* of mixed dialects⁵². It was subsequently used by Buddhist communities in western India before the oral teachings were transmitted to the monasteries of Sri Lanka during the Aśokan period (Oberlies 2019: 16-17). It is unclear when the Pali canon was first committed to writing, but this possibly occurred around the late first century BCE⁵³.

Although it is possible that discourse on the systematic knowledge of Prakrit may have occurred as early as the first century CE, available documents that reference grammatical characteristics of the literary Prakrits are only reliably dated from around the 10th century CE (Ollett 2017: 143-144, 205). Based on the available records of the Aśokan inscriptions and literary Prakrits, the grammars of the MIA dialects and Pali changed in significant ways compared with the known OIA languages, and these changes must have

⁵¹ The language spoken by the Buddha has been the subject of much academic discussion and controversy over many years. For some of the more recent contributions to this discussion refer to Karpik (2019a, 2019b), Levman (2019) and Wynne (2019).

⁵² E.g., Levman 2019; Norman 1992: 237; Oberlies 2019: 9.

⁵³ E.g., Salomon (2018: 56-57), Cousins (2019: 128) and Oberlies (2019: 5-6, 16-17). It has been suggested that some elements of Pali became Sanskritised, meaning that earlier OIA forms were gradually restored in the written canon, during and following its formation in the 1st century BCE up to around the 4th century CE, leading to a view that Pali is to some extent an artificial language. This view is contested by Karpik (2019a: 11-12), who argues that Pali is a natural standard language. For recent discussions on this issue refer to Cousins (2013), Karpik (2019a, 2019b) and Levman (2019), for example.

occurred at a much earlier time than the attested records. While there are variations in the nature and extent of the recorded changes, the following patterns are generally observed⁵⁴:

- Sound changes such as weakening of intervocalic stops, de-aspiration of aspirated stops, reduction of diphthongs to monophthongs, merger of the sibilants, assimilation of conjunct consonants as geminates or through epenthesis, loss of distinction between the semivowels.
- Breakdown of the complexity of the OIA case morphology in nominal forms, to the extent that this no longer served as a functional declension system in some languages.
- Erosion of the wide number of verbal forms which existed in the OIA languages.

While knowledge and analysis of Gāndhārī remains incomplete at the present time, it is assumed, similar to the other Prakrits, that it was not a direct descendant of either Ṛgvedic or Classical Sanskrit, but developed from another archaic OIA dialect, possibly pre-dating but similar to the Sanskrit of the Ṛgveda. As noted previously, Gāndhārī displays a mixture of advanced and conservative forms, and retained or developed a number of unique features, including retention of OIA *Tr* consonant clusters and non-merger of the three OIA sibilants, *ś*, *ṣ* and *s*. These were possibly influenced by the unique geographical position of Gandhāra in the northwest of the Indian sub-continent, and some of the sociolinguistic conditions peculiar to this region:

- The relative isolation from the rest of South Asia and the other developing MIA languages⁵⁵, which may have led to the retention of or slower development away from conservative OIA features. It is also possible that Gāndhārī is descended from the dialect of the first Indo-Aryan migrants to have settled in the area, while other groups moved further to the east where the MIA languages developed in different ways.
- The proximity of Gandhāra to the mountainous regions of the northwest, and the Dardic languages spoken in these areas (§ 2.2.6). As discussed in § 2.2.6, there are similarities between some of the consonant developments in Gāndhārī, such as retention of the three sibilants, and those preserved in the modern Dardic languages, suggesting possible areal influences between Gāndhārī and the more isolated northwestern languages from the earliest times.

⁵⁴ Masica 1991: 167-187, 230-237, 257; Oberlies 2019: 22-33.

⁵⁵ E.g., Salomon 1999: 3-5.

- The location of Gandhāra at the junction of major trading routes to the west through modern-day Afghanistan and Iran, as well as into the old Silk Road cities of Xinjiang in Central Asia, which would have brought many diverse language influences. Just as Gāndhārī was used as a vehicle for the transmission of Buddhism along these trading routes, it is quite likely that influences from some of the regional languages were transmitted back into Gandhāra through the agency of local traders.
- The retention of conservative characteristics could also be due to a stronger Sanskritic influence in one of the early centres of Brahmanic culture in the north west, and home of the Pāṇinian grammatical tradition⁵⁶. This may have manifested in some form of linguistic conservatism among the intellectual communities of the area, either in terms of retained, borrowed or even re-constructed Sanskritic forms in scholarly Gāndhārī.
- At the same time the movement of Buddhism into the area would have brought a strong influence of eastern features, especially in the copying of texts from manuscripts originally composed in other dialects, seen for example in the number of eastern forms found in some manuscripts.
- While this did not necessarily influence the structure of the language, the recording of Gāndhārī in the Kharoṣṭhī script affected the way in which the sound system was represented.

2.2.4 Languages or dialects?

Dialects are mutually intelligible varieties within the continuum of a language. Karpik (2019a: 15-19) has posited several pieces of evidence that point towards a mutual intelligibility between the MIA languages. These include the use of certain terms in the Buddhist Vinaya which suggest that “Aryan” was considered to be a single language by the Buddhist community of the time, along with the many similarities observed in the Aśokan inscriptions. And while many Buddhist terms have variations in meaning, this is not uncommon among mutually intelligible varieties, such as in some of the terminologies used by speakers of British and American English. This comparison points towards many other examples of a similar nature, and only touches on the wider subject of how divergent dialects need to be in phonology, morphology, and syntax before they become mutually unintelligible at a spoken level (Masica 1991: 23-27). Although the regional varieties of

⁵⁶ Burrow 1972: 35-36; Deshpande 1983: 110; Salomon 2018: 15; Oberlies 2019: 9-10.

the MIA language family are commonly referred to as dialects⁵⁷, there has also been much discussion about the possible existence of a MIA lingua franca or koine dialect, suggesting that some of the varieties were sufficiently distinct to help sustain the presence of such a dialect. In one approach it has been suggested that Pali may have developed into or played such a role⁵⁸, while an alternative interpretation suggests that this role was served by another similar but slightly different language, one possibly more aligned with the northwestern Prakrit of Gandhāra⁵⁹.

While a number of differences are recorded between the phonology and morphology of the Aśokan Prakrits, as noted by Karpik (2019a: 15-19), a significantly greater number of identical or very similar forms are shared between the dialects. This pattern can be observed in various Synoptic tables which compare the phonological and morphological developments in the inscriptional Prakrits⁶⁰, to the extent that the individual variations seem to be of lesser consequence in a wider comparison of lexical items. This is also seen in the comparative Synoptic texts of the Aśokan Prakrits⁶¹, where many words have similar or identical spellings, such as in the comparisons of *raja* / *raya* / *lājā*, or *samājā* / *samaya* / *samājā*, and *Devānampriyasa* / *Devanapiasa* / *Devanampiyasa* for example, where the words in each grouping reflect usage in the western, northwestern, and eastern versions respectively of the Aśokan inscriptions (Hultzsch 1925: 185). And while the variations reflect important divergences between the dialects, much commonality remains between the overall sound pattern of these words. This also extends to the overall sentence and clausal structures, and just as modern scholars can read these texts alongside each other without significant difficulties, it is most likely that an experienced translator, familiar with the dialects of the different regions, would also readily comprehend the intent of these texts, at least in a scholarly environment. This could however become a more awkward and labored exercise in the spoken language if the speakers are not fluent in each other's dialect, rendering normal communication more difficult in practical environments such as trade for example, and pointing towards an impetus for development of a simpler form of koine dialect across the region. It has been suggested that such a dialect could have derived from an earlier pre-canonical language which was the common

⁵⁷ E.g., Norman 1992: 237; Salomon 2018: 63; Oberlies 2019: 16.

⁵⁸ E.g., Norman 1992: 237; Oberlies 2019: 9, 16

⁵⁹ Levman 2014, 2016, 2019.

⁶⁰ E.g., Mehendale 1948: 203-268.

⁶¹ E.g., Hultzsch 1925: 185-213, Appendix I.

source from which later regional variations developed⁶². Hence, rather than being a created or merged dialect which developed due to the needs of increasing trade and religious contact, this koine was possibly a simpler and more levelled dialect emanating from an earlier stage of MIA development (Levman 2016: 32-36), whether this be Pali or another similar dialect. Such a koine would have sustained itself through the need for a common form as the regional dialects diverged from each other in geographically separated environments.

2.2.5 Source materials utilised by Gāndhārī scribes

The Aśokan inscriptions appear to have been translated into local dialects from Māgadhī, the central Magadha Prakrit of Aśoka which was used in the Dhauli and Jauguḍa inscriptions⁶³. Since these were intended as edicts or proclamations of expected behaviours directed towards regional populations⁶⁴, they presumably represent careful translations into the regional literary dialects of the time. Buddhism had been established for only one or two centuries when it was introduced into Gandhāra during the reign of Aśoka around the third century BCE, and the recording of Buddhist stories and teachings in written manuscript form is thought to have begun in the late first century BCE (§ 2.2.3). The earliest written Gandhāran Buddhist sources are short inscriptions of various types dating from around the same time, and subsequently in longer birch bark manuscripts dating from around the first century CE, and possibly even the late first century BCE (§ 3.4). These birch bark manuscripts preserve texts of a variety of genres of Buddhist literature, and many appear to be translations prepared from original manuscripts written in other MIA languages⁶⁵. While this is not unexpected given the importation of Buddhism into the area following its flowering in eastern India, some of the texts in the manuscripts do not appear to be associated with texts known from other Buddhist traditions, raising the question of whether such materials were possibly composed and recorded entirely within the Gandhāran tradition (e.g., Salomon 2018: 83-91). And although Buddhists only used vernacular languages in the early transmission of the teachings, a limited number of Buddhist writers may have used an early form of Buddhist Hybrid Sanskrit (BHS) from as early as the first century BCE. This was however still largely based on a MIA language

⁶² Levman 2014: 637; 2019: 74-76.

⁶³ Misra 1967: 102; Salomon 1998: 135-136; Oberlies 2007: 165-166. There are some variations to this pattern such as the Gīrnār inscription which appears to have been composed in a dialect closer to Pali.

⁶⁴ E.g., Hultzsch 1925: xlvii-xlviii; Salomon 2018: 19-20

⁶⁵ E.g., Salomon 1999: 11; 2018: 59-62; Allon 2021: 52-59.

which was only Sanskritised to a limited extent at this time⁶⁶, and it seems most unlikely that any of the source texts used by Gāndhārī scribes were in Sanskrit or BHS. A gradual ‘Sanskritisation’ of some elements of the language are apparent in later third century Gandhāran materials, but this appears to reflect the growing influence of Sanskritic intellectual culture on written Buddhist documents from around that time (§ 8.7).

While the administrative and Buddhist literary forms of Gāndhārī are distant from the Prakrit literary culture of northern India (e.g., Ollett 2017: 11-18), they were used by educated officials and Buddhist monks, reflecting a higher prestige variety of the regional language. But as noted by Houben (2018: 12) even the Vedic pundits may have spoken some form of Prakritic idiom when not engaged in ritual activities. Similarly, the Gandhāran scribes must have spoken some form of regional Gāndhārī dialect, and elements of this speech may be reflected in the source materials in different ways. For example, the variable, infilled parts of formulaic inscriptions are more likely to reflect real evolution of the language, as are apparently anomalous or abnormal forms, and when different forms of the same word present in the same text, the more advanced form is most likely closest to the actual use (Fussman 1989: 453 [§16.3]). Such patterns have also been observed in the language of some of the Buddhist manuscripts (Silverlock 2015: 215-216), while various avadāna and pūrvayoga texts represented in the manuscript collections relate popular stories in a free-flowing non-formulaic style, and may bear a closer relationship to the spoken language (§ 3.3).

2.2.6 Dardic languages

The term ‘Dardic’ is used to describe a group of IA languages of the Northwest of South Asia, which occur in the generally isolated mountainous regions of the Hindu Kush, Swat and Indus Kohistan, the Karakoram and Western Himalayas. They do not appear to have any common defining features, and although some archaic features are preserved to different degrees because of isolation from general MIA developments, some innovative features have also developed⁶⁷. Gandhāra is located in the more open valleys immediately to the south of, and partially in the foothills of these mountain ranges (refer to Map 1), and it has been suggested that Gāndhārī shares sufficient similarities with the Dardic languages

⁶⁶ Edgerton 1936: 503; 1953: 1 [§ 1.4], 4-5, [§§ 1.33, 1.34].

⁶⁷ Bashir 2003: 821-822; Masica 1991: 462.

to be aligned with that group⁶⁸, and more specifically with Torwali (e.g., Burrow 1937: vi). Torwali is spoken by about 60,000 people in the upper Swat valley, slightly beyond the main extent of the old Gandhāran monasteries and related art finds in the lower Swat valley⁶⁹, suggesting the possibility that it may have natural affinities with Gāndhārī, or even be a descendant language. Its linguistic structure, including the phonology, has been described to a reasonable extent by authors such as Grierson (1929), Lunsford (2001) and Bashir (2003). Baums (2009: 162-166, 172-179, 185) has discussed similarities between some of the consonant cluster developments in Gāndhārī and those preserved in other modern Dardic languages, such as Kashmiri for example. Some of the distinct features in Gāndhārī, such as retention of the sibilants for example, are features shared with the modern northwestern languages (Hock 2015: 122-123), suggesting an areal influence or interaction between Gāndhārī and the northwestern languages of the time. There is however a significant gap in the language record between the end of the Gāndhārī manuscript tradition around the fourth century CE and the first recording of the Dardic languages during the nineteenth century, making it difficult to trace and firmly establish any historical linguistic connections and mutual influences through the intervening fifteen centuries. While comparison of Gāndhārī with the wider group of Dardic languages is a complex task beyond the scope of this thesis, a comparison of similarities and differences between the Gāndhārī and Torwali phonologies has been included in § 8.2.6.3. The historical development of a language such as Torwali will reflect both its relative isolation, as well as any contact influences from other northwestern languages, and divergences from Gāndhārī as a possible ancestor language may reflect these influences.

2.3 A sociolinguistic perspective

Taken together, the changes which occurred in the MIA languages, including phonemic mergers and weakenings, breakdown of the nominal case system morphology, and reduction of the number of verbal forms (§ 2.2.3), can be read as amounting to an overall simplification of the grammar⁷⁰. While it is possible to document these changes at a typological level, the task of correlating them with particular motivations or causations

⁶⁸ Burrow 1937: vi; Fussman 1989: 439 [§3.1]

⁶⁹ Masica 1999: 21; Zwalf 1996: 11-16; Lunsford 2001: 1-5; Bashir 2003: 864-865.

⁷⁰ Linguistic complexity can be difficult to define and depends on many aspects of a language. Trudgill (2011: 20-22) identifies three factors which can be used to help identify simplification, including regularisation of irregularities, an increase in morphological transparency, and the loss of redundancies. At a phonological level, this might be considered in terms of loss of contrasts, such as through sound mergers for example (e.g., Trudgill 2011: 116-118).

for change, whether these be of a phonetic, phonological, or sociolinguistic nature, is more difficult. As attested by early MIA forms recorded in the Ṛgveda, many of these changes occurred at an early stage in the MIA development, following migrations of the PIA and OIA speakers into the northwest of South Asia, possibly from as early as around 1500 BCE. At the same time the OIA languages spoken and used by the Indo-Aryan social elites remained relatively stable, changing only slowly compared with Vedic Sanskrit, until codified by the ancient grammarians around 400 BCE. This effectively preserved, in Classical Sanskrit, the language used by the educated Brahmins of the northwest. The contrast between the MIA languages and Sanskrit highlights the difference between these two distinct sociolinguistic environments. In the one instance this involved the more extensive and vibrant trade and other social interactions within the sphere of the vernacular languages, including interactions between the indigenous populations and the migrating IA speakers, while the other involved the more controlled environment of a social and religious elite. The differences between these two environments, together with the overall simplification which occurred in the grammar of the MIA languages, suggest that sociolinguistic influences may have played a significant role in the development of the vernacular languages.

The extent to which correlations may exist between language and social structure, and in particular which influences, if any, may lead to the development of more complex or less complex grammatical structures, is an ongoing subject of scholarly study and debate. It has been proposed, for example, based on the study of a number of historical language developments, that languages tend to complexify or remain stable if they are undisturbed by contact, while languages tend to become simplified in certain types of contact situations⁷¹. It is shown, for example, that differences arise between dialects which exist in more exposed or high contact situations, compared with those existing in more isolated areas. In these cases, the more exposed dialects typically develop more innovative forms, while the more isolated ones retain conservative forms, reflecting a more archaic version of the language. Along with high levels of language contact, factors such as migrations into new territory or social upheavals all play a role in the disturbance of language stability, and the development of innovative forms appears to occur more rapidly

⁷¹ E.g., Trudgill 2002, 2011, McWhorter 2007. Simplification is more likely to occur in contact situations where a language is used as a lingua franca by large numbers of second language speakers over a long period. These observations are supported through many examples and more detailed analysis by both authors.

during such events⁷². In these cases, the innovative developments tend to arise as simplifications of the dialect. This is seen for example in the modern day Scandinavian languages descended from a common Old Norse ancestor, where the case morphology is significantly reduced in continental languages such as Danish, Swedish, and Norwegian, while the more isolated languages such as Icelandic and Faroese retain more conservative forms (Trudgill 2011: 3-4, 19-20). Such interpretations are not universally agreed however, and it has been argued, based on a number of counter examples, that a clear correlation between language complexity and the extent of language contact or isolation has not been established (e.g., Campbell 2013: 341-343). On the other hand, as suggested by Trudgill (2011: xv-xvi), these social factors are better interpreted as influencing factors, reflecting more common associations or tendencies, rather than being a sole determining factor for the changes. In this context, and as observed by Karpik (2019b: 114-115), interesting parallels exist between some of the examples of language simplification addressed by scholars such as Trudgill, and the simplification of the Prakrit languages associated with the Indo-Aryan migrations into South Asia. Middle English, for example, shows a marked simplification of inflectional forms compared with Old English. It is suggested this process of simplification developed in part through adult language learning among the socially subordinate majority of Celtic Late British speakers as they interacted with the socially dominant minority of Anglo-Saxon Old English speakers (Trudgill 2011: 50-55).

The Indo-Aryan migration into and across South Asia would have created many complex patterns of language contact and interchange with indigenous populations, and it is likely that a range of sociolinguistic factors similar to these known cases would have influenced the MIA language development. Differences between the sociolinguistic environment of Sanskrit, preserved within the more controlled surroundings of a social and religious elite, and that of the MIA Prakrits, which developed within the more extensive and vibrant trade and other social interactions of the vernacular languages, highlight the importance of understanding the Prakritic developments through a sociolinguistic perspective. However, beyond the limited lexical information recorded in the early Vedas, the absence of more detailed records of the early Prakritic developments in the northwest makes it very difficult to interpret these developments in any more detail, and nor is this within the scope of this thesis.

⁷² E.g., Trudgill 2011: 1-20.

3. Gāndhārī source materials

Modern knowledge of the Gāndhārī language only started in the late nineteenth century, when the Kharoṣṭhī script was deciphered (Salomon 2018: 68-72). Despite the recovery and translation of some important documents and manuscripts during the twentieth century, the extent of available source materials remained limited. But since the early 1990s a significant number of Buddhist manuscripts have been discovered, to the extent that a reasonable body of Gāndhārī source material is now available. Apart from the legends on coins minted by many rulers in Gandhāra, which are of limited linguistic value (Salomon 2019: 113)⁷³, the main Gāndhārī source materials are described below⁷⁴.

3.1 Inscriptions

Inscriptions in the Gāndhārī language and Kharoṣṭhī script fall into two main groupings:

- The third century BCE Aśokan rock inscriptions (Aś) at Shāhbāzgarhī and Mānsehrā in northern Pakistan (Hultzsch 1925: xi-xii, 50, 71), which provide the earliest significant record of the Gāndhārī language.
- A corpus of non-Aśokan Kharoṣṭhī inscriptions (InsK) of diverse types including many Buddhist rock inscriptions as well as inscriptions on seals, vases, sculpture pedestals, reliquaries and the like. While they are typically short and have a repetitive style and limited vocabulary (Salomon 2002: 121), they do provide a useable linguistic resource. A compilation and translation of 90 or so known non-Aśokan inscriptions was published by Konow (1929), while updated collations and more detailed descriptions of particular items have been published by a number of authors. These include for example Fussman (1989), Falk (2003, 2008) and Strauch (2012a) amongst others, while a new catalogue of the reliquary inscriptions, which includes editions and translations, has been published by Baums (2012). These inscriptions are generally located or were uncovered in northern Pakistan and eastern Afghanistan, and are dated between the first century BCE and second to third centuries CE.

⁷³ For descriptions of some of the typical coins minted under different rulers see Salomon 2018 (24-43).

⁷⁴ This section provides a general descriptive summary of the main Gāndhārī source materials. A more comprehensive itemised listing of the Gāndhārī source materials, including all individual items in each collection or grouping, is available through the online *Catalogue of Gāndhārī Texts* by Baums and Glass (2002b).

3.2 Administrative Documents

A significant body of administrative documents retrieved from Niya in the Xinjiang region of China, along with other sites including Lou-Lan, and sometimes referred to as the Niya Documents (ND). Recovered by Aurel Stein in the late 19th century, and more recently by Chinese archaeologists, they are written on wood and leather and demonstrate the use of Gāndhārī as an administrative language in the Shan-Shan or Kroraina kingdom during the third century CE. Although somewhat repetitive and limited in scope and vocabulary, they are nonetheless linguistically valuable. However, the language contains many Iranian loan words and Tocharian elements, which makes this form of Gāndhārī different to that of the northwest of the Indian subcontinent⁷⁵.

3.3 Buddhist manuscripts

The most significant collections of Buddhist manuscripts discovered and published or reported on to date include the following:

- A substantially complete copy of the Dharmapada, a well-known Buddhist literary text, written in Gāndhārī on a birch bark scroll. This was discovered as a complete manuscript near Khotan in Xinjiang during the 1890s, although exact details of the discovery remain unclear, and two parts were subsequently purchased by Europeans (Brough 1962: 1-2). Referred to as either the Khotan or the Gāndhārī Dharmapada (Dhp-G^K), it has been dated to around the first or second century CE (Brough 1962: 55). This document has been described in detail by Brough (1962), including discussion of the phonology and grammar. The form of Gāndhārī used in this manuscript shows systematic differences from that found in Gandhāran documents, suggesting that it was written in Central Asia (Falk 2015: 24, 52-59).
- A collection of twenty-nine fragments of birch bark scrolls, containing a variety of Buddhist religious texts, which came into the possession of the British Library in 1994. Referred to as the British Library Collection (BL), these scrolls had been stored in pottery jars, thought to have been found in the ruins of a monastery near Haḍḍa in eastern Afghanistan (Salomon 1999: xv, 15, 20-23)⁷⁶. They have been tentatively dated to around the first or second century CE, with some evidence suggesting as early as the

⁷⁵ Burrow 1937: v-vi; Salomon 1988; Salomon 1999: 113; Salomon 2002: 121.

⁷⁶ A detailed description of this find and its component parts is presented in Salomon (1999).

first century CE (Salomon 2000: xii). Although fragmentary, many of the texts contain substantially complete parts which are of significant linguistic value. The manuscripts appear to have been old and worn-out, and had been marked for ritual disposal as ‘dharma-relics’ (Salomon 2014: 2). Detailed analysis and translation of the text of nearly half of these fragments have since been published or are in preparation for publication (Salomon 2014: 4).

- A small collection of five scrolls reportedly found near the Pakistan-Afghanistan border area comprising the Mohmand Agency and Bajaur. These were originally sighted in 2005, and appear to have been part of a much larger find. Originally known as the ‘Split’ Collection (SC) for this reason, they have since been identified as part of the New Collection. Two of the manuscripts in the collection were radiocarbon dated, one in the first century BCE while the other is in the late first or early second century CE, which is more consistent with the stage of script development. The central date is 74 CE⁷⁷. All manuscripts in the collection have been published⁷⁸.
- A collection of fragments from about nineteen birch bark scrolls discovered in 1999 near Bajaur in northwest Pakistan. Known as the Bajaur Collection (BC), they appear to date from around the late first or early second century CE⁷⁹. While fragmentary, like most Gandhāran manuscripts, many of the texts appear to contain substantially complete parts, and detailed work has been undertaken on a number of the manuscripts, some of which have been published⁸⁰.
- A collection of twenty-four birch bark Buddhist scrolls and scroll fragments owned by the private collector Robert Senior in the United Kingdom. This collection is currently on loan to the University of Washington Library, and is referred to as the Robert Senior Collection (RS) (Salomon 2003: 74). Similar to the British Library Collection, these were stored in a pottery jar which was most likely also found near Haḍḍa, but in contrast to the British Library Collection, several of the manuscripts are complete or nearly complete. Based on radiocarbon dating of the manuscript and an historical era date inscribed on the pot lid, the collection has been dated to around 130-140 CE⁸¹. Inscriptions on the pot and its lid indicate that the manuscripts were commissioned and

⁷⁷ Falk 2011: 13; Falk and Strauch 2014: 54.

⁷⁸ Falk 2011; Falk and Karashima 2012, 2013; Falk 2015; Falk and Steinbrückner 2020, 2022.

⁷⁹ Strauch 2008a: 18; Strauch 2008b: 111; Falk and Strauch 2014: 51-54

⁸⁰ Schlosser and Strauch 2016a, 2016b; Schlosser 2016, 2022

⁸¹ Allon et al. 2006; Allon 2014: 22

newly prepared for ritual burial, presumably within a stūpa (Allon 2014: 22-23), in contrast to the disused manuscripts found in the British Library Collection. Detailed analysis and translation of the text of six of these scrolls have been published (Allon 2014: 19-21), either as part of the GBT series (2), as PhD theses (3), or in published articles (1), with further material forthcoming.

- Several hundred mostly very small palm leaf fragments written in Gāndhārī in the Kharoṣṭhī script, found near Bamiyan in Afghanistan, which are now divided amongst the Schøyen, Hirayama and Hayashidera Collections. Generally referred to as the Bamiyan manuscripts (BM), these appear to date from around the late second or early to mid-third century CE (Salomon 2014: 1, 6-7), and the orthography in some of the manuscripts displays later Sanskritic influences (Salomon 2001). A number of the texts had been identified among the fragments of the Schøyen Collection and the details of at least four of these have been published⁸².
- A small collection of eight fragments of palm leaf manuscripts held in the Bibliothèque National de France in Paris, known as the Pelliot Collection (PC). These were found near Kucha in northern Xinjiang in the early twentieth century by the French explorer Paul Pelliot. They have been difficult to date, but are possibly from around the same time or a little later than the Schøyen fragments (Salomon 2014: 8). They were published in full in Salomon (1998a), but are too limited and fragmentary to provide useable material.
- Another collection of manuscripts, reportedly found at an unknown location in Afghanistan and held privately for a number of years, was subsequently made available for inspection (Salomon 2014: 9-10). This collection, which is also associated with the Split Collection, was originally known as the New Collection (NC). It has subsequently been donated to the Islamabad Museum in 2023 and is now referred to as the Islamabad Museum Collection (IM) (M. Allon, personal communication, February 27, 2023). Most of the scrolls are in a fragmentary condition and are yet to be opened and conserved, although some scrolls have been conserved and studied, and preliminary details published⁸³.

⁸² Allon and Salomon 2000; Baums, Glass and Matsuda 2016; Baums, Braarvig, Lenz, Liland, Matsuda and Salomon 2016; Jantrasrisalai, Lenz, Qian and Salomon 2016.

⁸³ E.g., Harrison, Lenz, and Salomon 2018; Allon 2019.

A number of genres of Buddhist literature are represented among these manuscript collections, and an overview of this material is presented by Falk and Strauch (2014), including a classification and discussion of the texts belonging to various literary genres (p.57). These include a number of avadāna and pūrvayoga texts which relate popular stories based on local traditions, written in a language that may bear a closer relationship to the spoken language than other more formal texts. These are represented in the British Library, Pelliot and Split Collections (pp. 65-67).

3.4 Summary of source materials

A summary of these source materials is presented in Table 3.4, where they are tabulated according to their general area of origin and approximate historical era. Many of the particular locations in which these source materials originated or were found are identified on Map 1.

3.5 Selection of source materials for analysis

Fussman (1989: 439-440 [§ 3.3]) has suggested the existence of at least three types of Gāndhārī:

- The administrative jargon of the Niya documents, which also contains many Iranian loanwords.
- The literary Gāndhārī used in the Iranian countries and Central Asia which was in effect a foreign literary language.
- The literary and inscriptional Gāndhārī of northwestern Pakistan and eastern Afghanistan, which is in some way related to the living vernacular language in its home region.

The more recent discoveries of Buddhist manuscripts in the Greater Gandhāra region of northern Pakistan and eastern Afghanistan have significantly increased the scope of materials falling into this latter category, greatly expanding the ‘South Asian Corpus’ of Gāndhārī source materials. Because these originated or were found within the local area of the greater Gandhāra region it is more likely they are the work of scribes who spoke a version of the living vernacular Gāndhārī of their time in this region, and for traces of this to be evident in the documents. For this reason, the materials in this corpus are more appropriate for comparative study of the evolving language, and will be selected as source

materials for this thesis. These materials are recorded in the shaded cells in Table 3.4. In contrast, the Central Asian materials were found in an area that is geographically isolated and culturally different from Gandhāra. They are also written by people who were speakers of Iranian languages and not Gāndhārī, and the forms of Gāndhārī used vary through cultural influences such as Iranian loanwords.

Table 3.4: Locations, Types, and Historical Eras of Gāndhārī source materials

		Greater Gandhāra					Central Asia				
		Inscriptions		Manuscripts			Inscriptions		Manuscripts		
		in1	in2	m1	m2	m3	in1	in2	m1	m2	m3
BCE	3 rd c	Aś									
	2 nd c										
	1 st c	InsK	InsK								
CE	1 st c	InsK	InsK	BL SC & NC ⁽¹⁾					Dhp-G ^K		
	2 nd c	InsK	InsK	BC RS						PC	
	3 rd c	InsK	InsK		BM						ND
	4 th c										
Methods of Record Inscriptions in1 = inscriptions on rock in2 = other inscriptions Manuscripts m1 = birch bark m2 = palm leaf m3 = wood / leather Notes: (1): The SC and NC appear to be parts of a previously single collection. The NC is also now identified as the Islamabad Museum Collection (IM).					Collections: Aś = Aśokan Inscriptions BC = Bajaur Collection BL = British Library Collection BM = Schøyen and other Bamiyan Collections Dhp-G ^K = Khotan Dhammapada InsK = non-Aśokan Kharoṣṭhī inscriptions NC = New Collection ⁽¹⁾ ND = Niya Documents (Kroriana / Shan Shan) PC = Pelliot Collection (Kucha) RS = Robert Senior Collection SC = Split Collection ⁽¹⁾						

The available body of material from the South Asian corpus comprises the published descriptions of the Aśokan and other inscriptions, along with a significant number of manuscripts from the British Library, Robert Senior, Bajaur, Split, Bamiyan and New Collections. And while many more inscriptions and manuscripts have been transliterated, and in some cases are being prepared for publication, it is necessary to make a representative selection of the most appropriate materials from the published sources for analysis in this thesis. In selecting inscriptions for linguistic analysis Fussman (1989: 443-444 [§ 5]) suggested criteria of reliable location, dating, length, and assured interpretation. The locations and dating of the manuscript finds are discussed above, and most of the manuscripts selected for analysis and publication are relatively long. Confident interpretation of the texts requires that the Kharoṣṭhī script is at least in part clear and readable, and that Sanskrit equivalents of words can be found or presumed with reasonable levels of certainty. This process has been greatly enhanced by the presence in many situations of fully or partially parallel texts in other better-known Prakrit languages where the parallels to Sanskrit are better understood, including Pali, or in the Sanskrit language itself, including Buddhist (Hybrid) Sanskrit. Parallel texts are available in at least one of these languages for significant portions of all of the Buddhist manuscripts. The Aśokan rock edicts at Shāhbāzgarhī and Mānsehra are among a larger group of parallel edicts located across India and Pakistan, and all of the other edicts were written in Prakrit using the Brāhmī script, allowing comparative synoptic texts and glossaries to be prepared⁸⁴. Only limited parallel text is however available for the other inscriptions described by Konow (1929), Fussman (1989), and others, although Sanskrit equivalents are typically listed alongside the examples given in descriptions of phonological change. Another consideration in the selection of manuscripts is that where possible avadāna and pūrvayoga texts should be included in the selections. As noted above these may bear a closer relationship to the spoken language than other more formal texts.

The inscriptions and documents included in the published materials are described in more detail below, listed in the generally understood chronological order of the original materials. Most of these have been selected for use in this thesis, while others are not included for reasons discussed in the text.

Third century BCE:

⁸⁴ E.g., Woolner 1924; Hultzsch 1925: 185-213; Bloch 1950.

- The **Aśokan rock edicts** (Aś) inscribed at Shāhbāzgarhī and Mānsehra. These are described by Hultzsch (1925) in a general volume which covers, in considerable detail, all of the Aśokan inscriptions located across the Indian subcontinent, which were inscribed during a period around the approximate date of 250 BCE (Hultzsch 1925: p. xxxvi). The Hultzsch volume builds on the work of a significant number of earlier scholars including Cunningham (1877), Bühler (1889, 1892, 1894), Michelson (1909) and many others. It is also complimented by the parallel but less detailed work of Woolner (1924), which does however include a very useful glossary, along with that of Bloch (1950). Careful descriptions of the locations and features of each edict are provided, and plates of the rock faces showing the Kharoṣṭhī script, and Devanagārī and IAST transliterations and English translations for the Shāhbāzgarhī and Mānsehra inscriptions are included (Hultzsch 1925: pp. xi-xiii, 50-84). There is also a section on the grammar which includes a lengthy description of the ‘phonetics’ and lists some Sanskrit equivalents to Gāndhārī words (pp. lxxxiv-xc). The transliterated texts of the six major Aśokan rock edicts are presented in a comparative synopsis (pp. 185-213). Here the texts of the Shāhbāzgarhī and Mānsehra edicts are listed alongside the near identical texts of the Gīrnār and Kālsī edicts located in the West and North of India, and the Daulī and Jaugaḍa edicts located in the East, all of which were recorded in MIA Prakrits using the Brāhmī script (§ 2.2.3). A more detailed and well tabulated comparative study of the phonology of the Aśokan rock edicts is also provided by Mehendale (1948: 2-27, 203-237), building on the prior work of Hultzsch. These well documented inscriptions provide an important historical foundation for the analysis.

Second century BCE:

- No manuscripts and only a small number of inscriptions are attested.

First century BCE:

- Descriptions of the **non-Aśokan Kharoṣṭhī inscriptions** (InsK) of different types have been published by a number of authors:
 - Around ninety or so inscriptions which were known at the time of publication have been described and analysed in considerable detail by Konow (1929). Apart from some coins, seals and seal dies, almost all of the inscriptions are records of pious acts or donations of a Buddhist nature, made on a variety of materials including rock, stone, stone vases, pottery, silver scrolls and copper plates and ladles (p.

cxvi). The inscriptions have been dated generally by their connection to various historical eras, while some thirty-six of these bear actual dates within the particular era, which have been tentatively correlated with dates in the modern calendar (pp. lxxxii–xciv). These inscriptions span a period from the first century BCE through the first, second and third centuries CE. Similar to the work of Hultsch described above, Konow (1929) has provided a careful description of the locations and features of each inscription, along with plates of some of the material, although in places these are not very clear. IAST transliterations and English translations are also included, along with a grammatical sketch which includes a description of the phonology and lists Sanskrit equivalents of some Gāndhārī words (pp. xcv–cxi). A detailed and well tabulated presentation of the phonology, including Sanskrit equivalents, is also provided by Mehendale (1948: 295–314), which builds on the prior work of Konow and is usefully tabulated according to the progressive dates of the inscriptions. Thirteen of the inscriptions belong to the first century BCE (Mehendale 1948: 289–290).

- A significant number of other inscriptions, which became known between 1929 and 1989 and were published in diverse sources, have been collated by Fussman (1989). A detailed list of the inscriptions is provided, along with further selection of the most significant inscriptions in the group. Many of these inscriptions are Buddhist dedications, which include both a fixed portion of a formulaic and more archaic nature, as well as an infilled portion which may be closer to the actual usage and is potentially more revealing of the common usage of the time (Fussman 1989: 452–453 [§16.3]). One of these inscriptions belongs to the second century BCE, and four to the first century BCE (pp. 453–454 [§ 17.1]).
- A new catalogue of the Gāndhārī reliquary inscriptions which amalgamates some fifty-eight inscriptions described previously by different authors, and including updated editions and translations (Baums: 2012). Some forty-five of these have known or estimated dates, ranging from the second century BCE to the second century CE.
- Other publications which describe particular inscriptions or groups of inscriptions have been published by a number of authors including Falk, Fussman, and Salomon for example. While not presented as comprehensive compilations these are useful reference documents.

First century CE:

- Twenty-three of the InsK inscriptions described by Konow (Mehendale 1948: 290-291) and fourteen of those described by Fussman (1989: 453-454 [§ 17.1]) have been tentatively assigned to the first century CE.
- **British Library Collection** of Gāndhārī Buddhist manuscripts, comprising some twenty-nine birch bark scrolls or scroll fragments. These are tentatively treated as belonging to the first century CE based on the dating suggested by Salomon (2000: xii), although this evidence is not conclusive and these could also be second century manuscripts. The work of four different scribes is represented in the texts described in the materials published to date, known as BL scribes 1, 2, 4 and 9 (Salomon 1999: 54-55). The documents included in the published materials are described below, categorised according to the scribe which allows comparison of any consistency or divergence in the spelling habits of the scribes. Where available details of parallel texts in Pāli, Prakrit, Sanskrit or Chinese are provided in the publications, and the word lists in these publications typically include both Sanskrit and Pali equivalents, either drawn from the parallel texts, or on a 'presumptive' basis (e.g., Allon 2001: 325).

BL Scribe 1

- BL Fragment 1: Gāndhārī version of the Anavatapta-gāthā (AG-G^L) (GBT 5: Salomon 2008: 94). This is a long but incomplete scroll and although broken into five parts, these total about 160 cm in length with a typical width of about 14 cm (pp. 83-84), while the surviving text comprises some 126 lines (p. 87). This text provides a good representative sample of the orthography of Scribe 1, which is supplemented by the following documents.
- BL Fragments 12 and 14: Three Gāndhārī Ekottarikāgama-Type Sūtras (EĀ-G) (GBT 2: Allon 2001: 55)⁸⁵. These fragments are contiguous with a combined length of some 76 cm and a width of about 14 cm (p. 42), and constitute the lower part of a previously longer scroll (p. 4), while the surviving text comprises some 73 lines (pp. 46-47).
- BL Fragments 16 and 25: A New Version of the Gāndhārī Dharmapada (Dhp-G^L) (GBT 3: Lenz 2003: 6). These two fragments are of the same original manuscript

⁸⁵ A second text written by another scribe, comprising some brief outlines of avadāna type stories, is also contained on this manuscript (Allon 2001: 3).

(pp. 4-6). Only part of the Dhp-G^L is represented over about 15 lines in the hand of Scribe 1.

BL Scribe 2

- BL Fragments 16 and 25: A Collection of Previous-Birth Stories (PY-G) (GBT 3: Lenz 2003: 2). This text covers the major portion of the original manuscript formed by these two fragments, and was written by a different scribe to the Dhp-G^L. This is a pūrvayoga text with Sanskrit parallels to four of the pūrvayogas (p. 209).
- BL Fragments 1, 2, 3 and 21: Gandhāran Avadānas (Av-G) (GBT 6: Lenz 2010: 17). These avadāna stories are contained on a number of well-preserved scroll fragments of which Fragment 1 is quite extensive (p. 51) while the others are of a lesser extent (pp. 95, 105-107, 117). This group of texts provides a supplementary record of the orthography of Scribe 2.

BL Scribe 4

- BL Fragments 7, 9, 13 and 18: A Gāndhārī Commentary on Early Buddhist Verses (Nid-G) (PhD Dissertation: Baums 2009: 87). These fragments form part of what appear to be a manuscript comprising a number of single sheets (p. 67), and following careful reconstruction (pp. 57-67), substantial lengths of manuscript and text were recovered. These total some 223 cm in length and include 413 lines of preserved text (pp. 67-69). This text provides an excellent representative sample of the orthography of Scribe 4.

BL Scribe 9

- BL Fragments 5B: A Gāndhārī Version of the Rhinoceros Sutra (Khvs-G) (GBT 1: Salomon 2000: 58). This fragment has been recovered and re-assembled from twelve large sub fragments along with many smaller pieces. The re-assembled manuscript, although incomplete along its edges, contains a substantial amount of text (Plates 2 and 4, p. 110) and provides a good representative sample of the orthography of Scribe 9.
- **Split Collection** of Buddhist manuscripts. This collection comprises five birch bark scrolls or scroll fragments, in variable condition, and the nature, condition and content of the texts contained on the scrolls are described by Falk (2011). Each manuscript was

prepared by a separate scribe, and the numbering of SC 1 through to SC 5 is used to identify the five manuscripts, following the listing given by Falk (2011: 13). Based on radiocarbon dating of two of the manuscripts and the style of the language, this collection most probably belongs to the late first century CE (Falk and Strauch 2014: 54). Papers describing all five manuscripts in more detail have now been published, including comparative study of the Buddhist texts against similar Pali, Sanskrit and Gāndhārī renditions, along with an overview of the phonology and morphology of the language of the texts⁸⁶. These do not however include itemised description of the phonological changes like those presented in the longer GBT series and PhD publications of other manuscripts. At the same time, the texts appear to follow the typical sound change patterns of literary Gāndhārī, although some differences in the orthography have been reported⁸⁷. These manuscripts will not be included as primary source materials in this thesis for these reasons.

Second century CE:

- Thirty-four of the InsK inscriptions described by Konow (Mehendale 1948: 291-293) and two of the most significant inscriptions described by Fussman (1989: 453-454 [§ 17.1]) have been tentatively assigned to the second century CE.
- **Bajaur Collection** of Buddhist manuscripts. Paleographic studies suggest that these date in the first or second century CE, falling somewhere between the BL and RS collections in time (Strauch 2008a: 17-18). The collection consists of some eighteen manuscripts, which appear to have been written on by at least nineteen different scribes (Strauch 2008a: 8-13). The most comprehensive publications of manuscripts from this collection describe the texts of BC Fragments 4, 11 and 6⁸⁸. The manuscripts BC 4 and BC 11 both appear to have been written by BC scribe 5, and while the writing of BC 6 shows some similarities with this hand, there are also a number of small differences. These can however be interpreted as variations in the hand of scribe 5, which may be the result of a slightly more careful writing style or variations in the condition of the writing implement and surface (Schlosser 2022: 50-51). There do not appear to be any direct parallels to these texts, which appear to be scholastic

⁸⁶ Falk 2011: 13-15; Falk and Karashima 2012, 2013; Falk 2015; Falk and Steinbrückner 2020, 2022.

⁸⁷ Falk and Karashima 2013: 23-24; Falk 2015: 57-59; Falk and Steinbrückner 2020: 7-8, 13-14.

⁸⁸ The details of fragments 4 and 11 are described in Schlosser (2016), while all three fragments are described in Schlosser (2022).

documents such as the teaching notes of a *paṇḍita*, and were possibly created from memory or through dictation (Schlosser 2022: 21).

BC Scribe 5

- BC Fragment 4: Reconstruction of the surviving portions of this fragment yielded a scroll about 24 cm wide and 23 cm high, containing about 28 lines of text on the recto, and 12 lines of text on the verso. These appear to contain parts of what is possibly a proto-Mahāyāna text (Schlosser 2016: 26-28; Schlosser 2022: 10-11).
- BC Fragment 11: The surviving portions of this reconstructed fragment measure about 15.5 x 37.5 cm. It is written on both sides and contains some 81 lines of text altogether. These form part of a single scholastic text which appears to be some kind of commentary of the basic text on Fragment 4 (Schlosser 2016: 16-17, 30).
- BC Fragment 6: This fragment is in a poor state of preservation and appears to have measured approximately 30 x 10 cm in its original condition. It is written on both sides and contains about 20 lines of text and appears to provide further more specific commentary on the other two texts (Schlosser 2022: 4, 32). Although this manuscript only provides a limited sample of the orthography, it appears to contain some words that do not appear on fragments 4 and 11 (Schlosser 2022: 78-81).

While there is no Sanskrit or Pali parallel to the text on any of these fragments, detailed descriptions of the orthography and phonology are included, and presumed Sanskrit equivalents are given in the word index. Together these manuscripts provide a good representative sample of the orthography of BC Scribe 5.

- **Robert Senior Collection** of Buddhist manuscripts. As noted previously, the manuscripts and the pot in which they were found have been dated to around 130-140 CE. All of the manuscripts were prepared by a single scribe, and several of the manuscripts are complete or nearly complete. Detailed analyses of all or part of six of the manuscripts have been published. Where available details of parallel texts in Pali, Prakrit, Sanskrit or Chinese are provided, and Sanskrit and Pali equivalents are included in the word lists in all of these publications, drawn from either the parallel texts or on a presumptive basis (e.g., Silverlock 2015: 645).
- RS 5: Four Gāndhārī Saṃyuktāgama Sūtras (SĀ-G^{S1}) (GBT 4: Glass 2007). This manuscript consists of a full although partially damaged scroll, approximately 28

cm high and 27cm wide. It contains twenty-one lines of writing on each side, which make up three complete and one incomplete sutra (pp. 72-75). This text provides a good representative sample of the orthography of the RS manuscript scribe, which is supplemented by the following documents.

- RS 12: An Edition and Study of the Goṣiga-sutra, the Cow-Horn Discourse (Gos-G) (PhD Dissertation: Silverlock 2015). This manuscript forms half of a very long scroll, some 70 cm long and about 11 cm wide on average, and is one of the longest scrolls in the Senior Collection. It contains seventy-five (half) lines of verse, some fifty on the recto and twenty-five on the verso (pp. 126-130).
- RS 14: Gāndhārī version of the Anavatapta-gāthā (AG-G^S) (GBT 5: Salomon 2008). This manuscript is a partial remnant of half of a scroll that has been written on one side only. It is approximately 28 cm high and 10 cm wide, and contains twenty-five lines of writing (pp. 329-330).
- RS 19: A Study of the Gāndhārī version of the Dārukkhandhopamasutta (Discourse on the Simile of the Log) (SĀ-G^{S5}) (PhD Dissertation: Lee 2009). The scroll is complete and measures some 17 cm high and 21 cm wide (p. 3). It contains a single text of thirty-two lines, written on both sides.
- RS 20: Metaphor and Pedagogy in Early Buddhist Literature: An Edition and Study of Two Sūtras from the Senior Collection of Gāndhārī Manuscripts (SĀ-G^{S6}) (PhD Dissertation: Marino 2017⁸⁹). The scroll measures some 20 cm high and 22 cm wide (pp. 96-102). It contains a total of thirty lines of writing across both sides, which make up two sutras (pp. 72-75, 151-152).
- RS 22 nos. 2 and 3: A Gāndhārī *Samyukta-āgama* Version of the Discourse on Not-self (P *Anattalakkhaṇa-sutta*, Skt. **Anātmalakṣaṇa-sūtra* (Allon 2020); RS 22 no. 3: A Gāndhārī version of the Simile of the Turtle and the Hole in the Yoke (Allon 2007), which publishes only a portion of the sutra. These short sutras appear as the second and third of six texts written on scroll 22r, covering some twenty-three and twenty-six lines respectively of a very long scroll⁹⁰. The papers in which they are presented form a preliminary study, and although the edition text and Pāli,

⁸⁹ Updated versions of these Sūtras are also available online. Refer to Marino (2021a, 2021b).

⁹⁰ Allon 2020: 207; Allon 2007: 229

Chinese and Sanskrit⁹¹ parallels are presented, no detailed phonological analysis or word lists are available, and the papers are not utilised in this study.

While all the RS manuscripts were written by the same scribe, and the orthography, phonology and orthography are similar in all cases, the distinction between the individual manuscripts has been retained in the analysis. This allows a wider assessment of any variability in the recording of reflexes by the scribe, while also identifying this according to the manuscripts in question.

Third century CE:

- Twenty-three of the InsK inscriptions described by Konow (Mehendale 1948: 293-294) have been tentatively assigned to the third century CE.
- **Bamiyan Collection** of Buddhist manuscripts. This collection of many small palm leaf fragments, written in Gāndhārī in the Kharoṣṭhī script, was found near Bamiyan in Afghanistan. These have been dated to around the early third century CE, and at least forty-seven different scribal hands have been identified (Salomon 2014: 6). Fragments of known Buddhist texts have been identified on at least four sets of fragments, and descriptions of these have been published. These include Gāndhārī versions of the *Mahāparinirvāṇasūtra*⁹² (identified as the MPS-G^B)⁹³, the *Bodhisattvapiṭakasūtra*⁹⁴ (Bspt-G^B), and the *Bhadrakalpikasūtra*⁹⁵ (Bhk-G^B), along with fragments of an *Ekottarikāgama* manuscript⁹⁶ (EĀ-G^B). Although only fragmentary, the text of these manuscripts appears to represent a later stage in the evolution of Gāndhārī, involving a number of adjustments to the phonology and morphology of the language. Although not applied consistently throughout the texts, these restore some conservative Sanskrit features, in a process which has been described as “Sanskritisation” (e.g., Salomon 2001: 243). Due to its limited and fragmentary nature this material has not been included in the main analysis sections, but the types of changes observed are discussed further in § 8.7.

⁹¹ The Sanskrit parallel relates to the RS 22 no 2, while the Chinese parallel relates to the RS 22 no 3.

⁹² Allon and Salomon 2000.

⁹³ The B superscript indicates Bamiyan.

⁹⁴ Baums, Braarvig, Lenz, Liland, Matsuda and Salomon 2016.

⁹⁵ Baums, Glass and Matsuda 2016.

⁹⁶ Jantrasrisalai, Lenz, Qian and Salomon 2016.

3.6 Selected source materials

In order to facilitate the tabulation of data from the source materials, it is necessary to develop abbreviated ‘tags’ which identify the particular publication or manuscript within a collection, and also the scribal hand where this is relevant. This is more straightforward in cases where there is only one scribe, or where individual scribes cannot be distinguished on that basis, such for the Aśokan and non-Aśokan inscriptions. Here the sources can be identified according to the collection and the author of each reference publication, as for example in Aś-H and Aś-M for publications describing the Aśokan inscriptions prepared by Hultzsch and Mehendale respectively. Similarly, InsK-K, InsK-M and InsK-F identify publications describing the non-Aśokan inscriptions prepared by Konow, Mehendale and Fussman. Similarly, for texts in the RS collection which were all written by one scribe and where all of the publications to date deal with single manuscripts, these can be identified according to the manuscript number, as in RS-5, RS-12 and so forth. The situation is more complex in the case of the BL, SC, BC and BM collections which contain the work of multiple scribes, and where a number of similar manuscripts are often described in a single publication. In these cases, it is necessary to identify both the scribal hand as well as the reference publication. Hence the abbreviation BLS1 identifies a manuscript or group of manuscripts prepared by BL scribe 1, while BLS1-1 is one manuscript written by him, published in Salomon (2008), identified by the -1 suffix, while BLS1-2 is a second manuscript written by Scribe 1, published in Allon (2001), identified by the -2 suffix, and so forth.

The selected source materials are listed in their assumed historical order in Table 3.6, along with the ‘tags’ used to identify them. Where known, the individual scribes are listed separately for each collection, while the particular scrolls or scroll fragments from the collections which are studied in each publication are identified under the ‘Source’ column. These materials offer a reasonably comprehensive record of the developing Gāndhārī inscriptional and Buddhist literary language of the early centuries BCE and CE.

Table 3.6: Summary and identification of source materials – third century BCE to third century CE

Era		Coll. (1)	Scribe	Source material		
				Source (2)	Publication	Abbr. (3)
BCE	3 rd c	Aś	n/a	Aś	Hultzsch (1925)	Aś-H
					Mehendale (1948)	Aś-M

BCE	1 st c	InsK	n/a	InsK	Konow (1929)	InsK-K	
					Mehendale (1948)	InsK-M	
					Fussman (1989)	InsK-F	
CE	1 st c	InsK	n/a	InsK	Konow (1929)	InsK-K	
					Mehendale (1948)	InsK-M	
					Fussman (1989)	InsK-F	
		BL	1	BL 1	Salomon (2008) (GBT 5)	BLS1-1	
				BL 12, 14	Allon (2001) (GBT 2)	BLS1-2	
				BL 16, 25	Lenz (2003) (GBT 3)	BLS1-3	
			2	BL 16, 25	Lenz (2003) (GBT 3) (pūrvayogas)	BLS2-1	
				BL 1, 2, 3, 21	Lenz (2010) (GBT 6) (avadānas)	BLS2-2	
			4	BL 7, 9, 13, 18	Baums (2009) (PhD)	BLS4-1	
		9	BL 5B	Salomon (2000) (GBT 1)	BLS9-1		
		2 nd c	InsK	n/a	InsK	Konow (1929)	InsK-K
						Mehendale (1948)	InsK-M
Fussman (1989)	InsK-F						
BC	5	BC 4 & 11	Schlosser (2016) (PhD)	BCS5-1			
			Schlosser (2022) (GBT 7)	BCS5-2			
		BC 6	Schlosser (2022) (GBT 7)	BCS5-2			
CE	2 nd c	RS	5	RS 5	Glass (2007) (GBT 4)	RS-5	
				RS 12	Silverlock (2015) (PhD)	RS-12	
				RS 14	Salomon (2008) (GBT 5)	RS-14	
				RS 19	Lee, Mei-huang (2009) (PhD)	RS-19	
				RS 20	Marino (2017) (PhD)	RS-20	
	3 rd c	InsK	n/a	InsK	Konow (1929)	InsK-K	
					Mehendale (1948)	InsK-M	
	Notes: (1): Coll. = collection (2): For manuscript sources the scroll or scroll fragment numbers from the collection which are studied in the publication are indicated. (3): Abbr. = abbreviated ‘tags’ for sources				Collections Aś = Aśokan Inscriptions BC = Bajaur Collection BL = British Library Collection InsK= non-Aśokan Kharoṣṭhī inscriptions RS = Robert Senior Collection		

4. Sanskrit phonetics and phonology

Like the other MIA languages, Gāndhārī descended most directly from an early OIA dialect, which was likely similar to but not identical with Vedic Sanskrit (§§ 2.2.2, 2.2.3). Despite these differences, the available records of both Vedic and Classical Sanskrit provide known reference points of similar languages against which to compare the MIA languages and Gāndhārī. Some of the phonological features of Sanskrit, which will assist in understanding the developments which occurred in Gāndhārī, are outlined in this section. While much of this material will be known to those versed in Sanskrit, an overview of various linguistic aspects is provided as a background to understanding the changes which occurred in Gāndhārī in the context of this thesis. Further detail of some of the material has been included in Appendix A.

4.1 The Sanskrit phonological system

The Sanskrit phonological system is described in ancient grammatical works such as Pāṇini's *Aṣṭādhyāyī* and phonetic works known as the *Prātiśākhya*⁹⁷. Although it has been recorded using a variety of scripts, in northern India Sanskrit is most commonly written in the Devanāgarī script, which has been in use since the early second millennium CE⁹⁸. This is based on an abugida system, in which each consonant hosts the neutral vowel 'a', with other vowels indicated by diacritic marks, while the vowels are also represented by their own characters when not preceded by a consonant. The absence of vowels in word final consonants is also marked with a diacritic. The alphabet is arranged according to the so-called *varṇamālā* sequence based on the places and manners of articulation, providing a clear guide to the articulatory phonetics of the alphabet⁹⁹. While early Brāhmī scripts failed to record some of the vowels, such as diphthongs and vocalic *r* (Strauch 2012: 133-136), these are all recorded in Devanāgarī. The Sanskrit phonological system is presented in Table 4.1, with the consonants organised in a similar format to the IPA, with the labials to the left and velars to the right in the table, and with the fricatives placed before the semi-vowels¹⁰⁰.

⁹⁷ Allen 1953: 4-7; Cardona 2007: 105-108.

⁹⁸ Salomon 2003: 98-99; Cardona 2007: 156.

⁹⁹ E.g., Allen 1953: 7-20; Strauch 2012: 134-135.

¹⁰⁰ This differs from the normal Sanskrit and MIA convention of ordering the velars first, either on the left of a table or the top of a listing, and the labials last, either on the right of a table or at the bottom of a listing, while the semi-vowels are typically listed ahead of the fricatives (e.g., Allen 1953: 20; Burrow 1972: 67).

Table 4.1: The Sanskrit Phonological System

		Labial		Dental		Retroflex		Palatal		Velar		Glottal	
		unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced
Stops	unaspirated	<i>pa</i>	<i>ba</i>	<i>ta</i>	<i>da</i>	<i>ṭa</i>	<i>ḍa</i>	<i>ca</i>	<i>ja</i>	<i>ka</i>	<i>ga</i>		
	IPA	p	b	t̪	d̪	ʈ	ፈ	c	j	k	g		
	aspirated	<i>pha</i>	<i>bha</i>	<i>tha</i>	<i>dha</i>	<i>ṭha</i>	<i>ḍha</i>	<i>cha</i>	<i>jha</i>	<i>kha</i>	<i>gha</i>		
	IPA	p^h	b^h	t̪^h	d̪^h	ʈ^h	ፈ^h	c^h	j^h	k^h	g^h		
Nasals			<i>ma</i>		<i>na</i>		<i>ṇa</i>		<i>ña</i>		<i>ṇa</i>		
	IPA		m		ɳ		ɳ		ɲ		ɳ		
Fricatives				<i>sa</i>		<i>ṣa</i>		<i>śa</i>				<i>-ḥ</i>	<i>ha</i>
	IPA			s̪		ʂ		ʃ				h	ɦ
Semi-vowels	tap						<i>ra</i>						
	IPA						r (alveolar)						
	approximants		<i>va</i>		<i>la</i>				<i>ya</i>				
	IPA		v (Labiodental)		l				j				
Vowels	short	<i>a</i>	<i>i</i>	<i>u</i>	<i>ṛ</i>	<i>ḷ</i>							
	IPA	ə	i	u	ɾ	ɭ							
	long and diphthongs	<i>ā</i>	<i>ī</i>	<i>ū</i>	<i>ṝ</i> (uncommon)	<i>e</i>	<i>ai</i>	<i>o</i>	<i>au</i>				
	IPA	a:	i:	u:	ɾ:	e	ai	o	au				
Other	Anusvāra	<i>ṁ</i> or <i>ṁ</i>	An unmodified nasal sound following a vowel										
	Visarga	<i>ḥ</i>	A post-vocalic aspiration generally used in word final positions										

Source: Compiled from Macdonell (1927: 2-6); Allen (1953:20), Burrow (1972: 67); Coulson (1976: 4); Cardona (2007: 156).

Each phoneme has been represented using the IAST characters, and the assumed IPA representation of the pronunciation is also given, drawing from the works of Allen (1953) and Coulson (1976). These are based on articulatory descriptions given in the ancient texts along with contemporary pronunciation guided by the continuous oral traditions of Classical and Vedic Sanskrit. While the semivowel *r* is classified phonologically as retroflex, the pronunciation is generally described as alveolar, that is midway between dental and retroflex, in the ancient *Prātiśākhya*s (Allen 1953: 53-55). This is similar to the practice in modern Sanskrit and the present-day NIA languages, which can be seen in X-ray tracings of the tongue position for Hindi retroflex *ṛ* (Ladefoged and Maddieson 1996: p. 27, Figure 2.11). The phonetic representations given in Table 4.1 must be regarded as approximations, however, as even with the benefit of historical documents and oral traditions, it is not possible to capture the precise phonological characteristics of a 2,500 year-old language.

4.2 The Sanskrit sandhi system

The Sanskrit sandhi system comprises a series of rules that seek to capture the spreading or assimilation of sounds across ‘junction’ points which arise in the formations of words and sentences. These rules are applied word internally between morphemes as well as externally between words in sentences and compounds, where words are often joined in sandhi. They have been applied almost universally in the extensive body of written Sanskrit scriptures and texts, so that the rules must be interpreted in reverse in seeking to understand the structure of any sentence or word. The main principles of the sandhi system are assimilation and the avoidance of hiatus, or ease of utterance. These patterns were analysed by the ancient grammarians, and the sandhi rules set down by Pāṇini in the *Aṣṭādhyāyī*, which offers a detailed description of what is considered to be the correct pronunciation of Sanskrit scriptures and texts¹⁰¹. A level of familiarity with this system is helpful when analysing developments in the MIA languages in comparison with recorded Sanskrit parallels, and a brief overview is presented in this section.

The rules of sandhi are comprehensive and describe all possible junctions that occur between vowels, between vowels and consonants, and between consonants. The rules are normally grouped under the headings of internal and external sandhi. Internal

¹⁰¹ E.g., Whitney 1924: 34-87; Macdonell 1927: 10-32; Oberlies 2003a: 1-51; Cardona 2007: 110-121.

sandhi is typically applied between morphemes, including suffixes and declensional endings, and while external sandhi is typically applied between words its rules are also applied across some internal junctures. External consonant sandhi is mostly dominated by regressive assimilation, as seen in the following examples:

- The spread of place in *tat + jayati > tajjayati*
- The spread of place and manner in *tat + śarīram > taccharīram*
- The spread of voicing in *ut + gama- > udgama-*, but *ut + tara- > uttara-*

The rules of internal sandhi are of most interest in understanding the interaction between changes to consonants and word morphology that developed in a language like Gāndhārī. While some of these are reasonably simple, others such as those relating to the spread of retroflexion between consonants are more complex. Retroflex consonants were not present in the reconstructed PIE language (Clackson 2007: 34), and arose from a number of influences during the development of OIA from PIr, including borrowings, analogical spread and assimilation. The internal sandhi rules which describe the spread of retroflexion from vowels to consonants and between consonants through assimilation are often compounded, seen for example when the preverb *ni-* is added to *sanna-*, the past participle of the verbal root $\sqrt{\text{sad}}$ (MW s.v. “sit down”) to give *ni-sanna*. This is typically lexified as *niṣaṇṇa-*, recording a two-stage spread of retroflexion according to the rules of internal sandhi – firstly to *niṣanna* due to *i* preceding *s*, and subsequently to *-ṇṇ-* following the development of *ṣ*. On the other hand, **nisanna* is not recorded in the language and no minimal pair can be found in this, and many other cases, which supports a contrastive distribution between *s* / *ṣ* and *n* / *ṇ* as distinct phonemes in this word. The development of retroflex consonants in OIA is a complex topic, and an overview of these developments, including a summary of the sandhi rules describing the spread of retroflexion, are discussed further in Appendix A1.

From a linguistic perspective, one consequence of sandhi is that the Sanskrit orthography records some of the allophones of particular phonemes, but using the graphemes for other phonemes to represent these. Hence in the above example the graphs for *ṣ* and *ṇ* are used to represent speech variants of the phonemes *s* and *n* in the environment of a preceding *i*. At the same time, no distinction is made between other realisations of these phonemes which are not conditioned in the same way, and may contrast in minimal pairs, and cases such as this where the graphs of these phonemes are

used to represent particular conditioned realisations of other phonemes. The wide application of this system, especially in the use of internal sandhi at the word level, makes the identification and analysis of contrastive pairs quite difficult. It is assumed that the initial impetus of the ancient grammarians in establishing the sandhi system was in the nature of phonetic enquiry, focused on understanding the actual sounds of the language in various phonetic environments. As these patterns were understood and formalised, they were described in sūtras, orally at first and later in written forms. At the same time, the allophones were only recorded using the recognised phonemic orthography of the *varṇamālā* alphabet, although finer levels of phonetic variation undoubtedly occurred at both local and regional levels. It is uncertain, however, whether this system was initially intended to be prescriptive, or whether this status was slowly accreted in time after the recording of the sandhi rules. As observed by Whitney (1924):

“This euphonic interdependence of the words of a sentence, which is unknown to any other language in anything like the same degree, is shown to be at least in considerable measure artificial, implying an erection into necessary and inevitable rules of what in the living language were only optional practices, by the evidence of the older dialect of the Vedas and the younger Prakritic dialects, in both of which these rules (especially as regards hiatus), are very often violated.” (Whitney 1924: p 35 [§ 101a]).

4.3 Sanskrit nasal consonants

Five nasal consonants, *m*, *n*, *ṇ*, *ñ* and *ṅ*, are recorded in Sanskrit, along with the *anusvāra* ‘pure nasal’ *ṁ* (Table 4.1). Of these only the labial *m* and dental *n* appear independently in word initial, medial and final positions, while retroflex *ṇ*, palatal *ñ* and velar *ṅ* never occur word initially and are mostly contingent on adjacent sounds (Macdonell 1916: 8-9). An understanding of the status of these three nasals in OIA, along with *anusvāra*, is helpful when analysing the Gāndhārī developments.

The retroflex nasal *ṇ* attains phonemic status at a minor level in OIA (Appendix A1), while the palatal and velar nasals *ñ* and *ṅ* are essentially non-contrastive allophones of *n*, used almost entirely in homorganic clusters with stops of their own class. Their use is seen for example in *pañca-* (MW, s.v. adj. “five”), *añjana-* (n. “ointment”), *aṅka-* (m. “a hook”) and *aṅga-* (n. “limb”), where they have assimilated the place features of the

following stop. Similar homorganic assimilation also occurs following the palatal stop in *jñāna* (n. “knowledge”), but such assimilation does not appear to occur following the velar stops, seen for example in *gnā-* (f. “goddess”). Place assimilation of nasals within nasal plus stop sequences is common in the languages of the world (Ladefoged & Maddieson 1996: 119-123), where the transition from nasal to stop is simply realised by a lifting of the velum. Although such sequences are sometimes referred to as pre-nasalised stops, and recorded as *–^mb-* for example with the suggestion of a unitary segment, there appears to be little phonetic distinction between these and normal nasal + stop sequences which are recorded as *–mb-*, although there is a prosodic difference between a unitary segment and a cluster which will divide as coda-onset.

Traditionally, *anusvāra* occurred in post-vocalic positions prior to certain consonants where it was described as a ‘pure nasal’, generated by the postvocalic breath passing through the nose and unmodified by any consonant. In Vedic it was influenced by following consonants, leading to nasalised glides, or perhaps to a nasalised transition between vowel and consonant articulation. In a further development *anusvāra* was used to mark word final nasals preceding a stop, usually *m* but also *n*, where homorganic nasals had been recorded previously. This was later extended to morpheme boundaries and even intra-morpheme positions within a word. This latter usage does not represent nasalisation in the original sense, but is a form of orthographic abbreviation that became quite widespread¹⁰². It is normally represented in Devanāgarī script as a dot above the preceding akṣara, which is transliterated as either *ṁ* or *m̐*. The nature of *anusvāra* is discussed further in Appendix A2.

4.4 Sanskrit consonant clusters

A tabulation of Sanskrit consonant clusters is presented in Appendix A3, providing a basis for analysis of the changes that occurred between the OIA and Gāndhārī clusters (§ 8.3). There are approximately 277 two consonant clusters, and a further 71 three consonant clusters, which are mostly formed as extensions of the two consonant cluster with the addition of a liquid or glide¹⁰³. All of the two consonant clusters occur word medially, while some 81 of these also occur word initially. It is necessary to consider two

¹⁰² E.g., Macdonell 1916: 10; Whitney 1924: 70-73 [§§ 70-73]; Allen 1953: 40-46.

¹⁰³ There are also longer clusters such as in *tārṣṇya* which has five consonants in a row, but these are less common.

important parameters in the understanding and analysis of consonant clusters, which are the relative sonority and the relative strength of the consonants. The concept of sonority is based on a number of factors including audibility and the extent of opening of the oral stricture, which generally give rise to a more energetic and sonorant acoustic signal, but has been difficult to define precisely¹⁰⁴. Sonority is an important parameter in the analysis of syllable structure and the changes to the syllable structure which arose in MIA languages such as Gāndhārī (§ 8.4). A common version of the sonority hierarchy is shown in Table 4.4.1¹⁰⁵, and similar versions have also been applied to Sanskrit consonants¹⁰⁶.

Hierarchies of consonant strength are typically defined in terms of the degree of oral stricture alone, and follow a slightly different sequence than the sonority hierarchy due to the focus on oral stricture. Strength hierarchies have been used in the analysis of consonant cluster assimilation based on the respective manners of the consonants, as for example in the analysis of Pali assimilation (§§ 6.1.3, 9.3).

Table 4.4.1: Scale of relative sonority – from least at top to most sonorant at bottom

Consonant manner	OIA consonants
Voiceless stops	<i>k, kh, c, ch, ʈ, ʈh, t, th, p, ph</i>
Voiced stops	<i>g, gh, j, jh, ɖ, ɖh, d, dh, b, bh</i>
Voiceless fricatives	<i>s, ʃ, ś, h</i>
Voiced fricatives	
Nasals	<i>ṇ, ñ, ṇ, n, m</i>
Liquids	<i>r, l</i>
Glides	<i>y, v</i>
High then low vowels	

Source: Davenport and Hannahs (2010: 75)

4.5 Sanskrit syllable structure

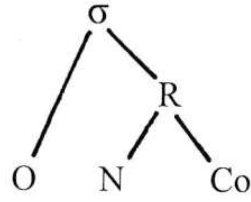
While there does not appear to be a clearly understood phonetic basis for syllables, they form part of the speech patterns that are recognised by most speakers of a language, and also play an important role in many linguistic processes and rules. Linguistic analysis typically separates the syllable (σ) into an onset (O) and rhyme (R), which further

¹⁰⁴ Clements 1990; Wright 2004: 50-51.

¹⁰⁵ E.g., Davenport and Hannahs 2010: 75.

¹⁰⁶ E.g., Cho 1999: 159; Kobayashi 2004: 59.

separates into a nucleus (N) and coda (Co), as in $\sigma = O + (R = N + Co)$, where the nucleus is the only obligatory element¹⁰⁷. This is shown in the following diagram:



The nucleus is generally the most sonorous sound, which is typically a vowel, and the patterning of syllables is built up around the sequence of vowel sounds and intervening consonants. The onset and coda can be made up of one or more consonants, and although the ways in which consonants are assigned between codas and onsets at the syllable margins vary in some languages, sonority sequencing provide good indication in most cases. This can be expressed in terms of the Sonority Sequencing Principle (SSP), which states that “*Between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted*” (Clements 1990: 285). This means that consonant clusters with increasing sonority are typically allowed as syllable onsets, while only clusters with falling sonority can reside in the coda, although sonority plateaus and reversals have been observed in some languages (Gordon 2016: 98-100)¹⁰⁸. The single syllable English word ‘crank’ provides a good example of this structure, with the cr- onset having rising sonority prior to the vowel nucleus -a-, which is followed by the -nk coda having falling sonority. (Davenport and Hannahs 2010: 74-75). Furthermore, the arrangement of segments at syllable boundaries mostly follows the Syllable Contact Law (SCL), which indicates that when two consonants meet at a syllable boundary, there is a preference for the first one to be more sonorant, or in other words, for sonority to continue to fall across the boundary (Clements 1990: 286-287). Hence clusters with rising sonority are typically assigned as onsets to the following syllable, with the exception that they must also be possible as word initial clusters in the language, as in Sanskrit {#, .} *tr-* for example. This means that clusters with rising sonority which do not occur word initially will be assigned as coda-onset pairs which divide across the syllable boundary, creating an exception to the SCL, as in Sanskrit *-b.l-* for example which does not occur word initially. On the other hand, clusters with falling sonority do not occur word initially¹⁰⁹, and generally form coda-onset

¹⁰⁷ Davenport and Hannahs 2010: 73-75; Gordon 2016: 83-84.

¹⁰⁸ A common exception to these principles occurs in relation to *st-* and *-st-* clusters, which are discussed further in Appendix A4.

¹⁰⁹ Apart from exceptions such as the *st-* clusters.

pairs across the syllable boundary in medial positions in accordance with the SCL, as in Sanskrit *-l.b-* for example. This can also be seen in the English word *cran.ky* where the *-nk-* grouping with falling sonority divides across what is now a syllable boundary. While generally applicable, there are issues and exceptions with both of these principles which relate to definitions of sonority as well as language permissible onset structures¹¹⁰.

According to ancient grammatical treatises the vowel forms the nucleus of the syllable in Sanskrit. Initial and medial single consonants, as well as word initial clusters, are all attached as onsets to the following vowel, while rules for the distribution of consonants in medial clusters appear to vary depending upon the dialect¹¹¹. Some texts suggest that the first consonant in medial clusters should be attached as a coda to the preceding syllable, while others suggest this is optional and they may be either divided as a coda-onset pair, or attached in full as onsets to the following syllable. More specific rules are given in another text, which suggests that clusters of consonants plus semivowel, or of stop plus fricative, be attached to the following syllable. This recognises the patterns of rising sonority in these types of clusters, which are very common and occur as word initial clusters in many cases¹¹². And while this word initial condition does not appear to be stated in the ancient texts, it is supported by reference to ancient Greek and Latin texts (Allen 1953: 82-83). In the assessment of syllable lengths in poetic prosody, which is based on sequences of short and long, or “light” and “heavy” syllables, medial clusters typically contribute length to preceding syllables, as for example where ‘heavy’ syllables include those with a short vowel followed by a consonant group (Allen 1953: 83-86). This requirement is satisfied by those clusters which divide as coda and onset, but may contribute to doubling of consonants across syllable boundaries in clusters having rising sonority. This is seen in the example of *yajñásya* > *yaj.jñás.sya*, where the coda *j* of the first syllable preserves syllable weight, while the *jñ* onset in the second syllable reflects resyllabification and respects sonority sequencing, and similarly with the following doubling of *s.sy*. This points towards a different assessment of the syllabic patterns of consonant clusters based on metrical requirements rather than one based purely on sonority. The syllable patterns formed by the Sanskrit consonant clusters are described in more detail in Appendix A4, while Sanskrit doubling phenomena are discussed in § 4.6 and explored in more detail in § 9.3.3.

¹¹⁰ Clements 1990: 287-290; Wright 2004: 34-35; Gordon 2016: 100.

¹¹¹ Allen 1953: 81-82; Kobayashi 2004: 39-41.

¹¹² Refer Table A4.1 in Appendix A4.

4.6 Sanskrit geminate consonants

Geminated or phonetically long consonants are found in many languages, and can form length contrasts with short consonants (Ladefoged and Maddieson 1996: 91-94; Gordon 2016: 51-57). Geminate consonant clusters are recorded in medial positions in both Vedic and Classical Sanskrit for all manners and places of articulation, including *-pp-*, *-bb-*, *-tt-*, *-dd-*, *-ṭṭ-*, *-cc-*, *-cch-*, *-jj-*, *-kk-*, *-mm-*, *-nn-*, *-ṇṇ-*, *-ss-*, *-ṣṣ-*, *-śś-*, *-vv-*, *-ll-* and *-yy-*¹¹³. The etymology of these geminates is variable and complex and can only be addressed here in general terms. One common synchronic source of phonological geminates in OIA is morphological, where they can develop through the addition of consonant final or initial affixes to consonant initial or final stems respectively. This is seen for example in Sanskrit *ut-tara* (MW. s.v. cpv. “upper, higher, superior”) which is formed through a common morphological process involving the prefix *ut-*, and in past participles such as *cit-ta* (MW. s.v. pp. √*cit*, “noticed”) and *yat-ta* (pp. √*yat*, “endeavored”). While the geminate *-tt-* is not contrastive in *uttara*, whose potential minimal pair of **utara-* is not listed in MW. for example, in *citta* and *yatta* it is contrastive with short *-t-* in other past participles such as *ci-ta* (pp. √*ci*, “heaped”) and *ya-ta* (pp. √*yam*, “restrained”) for example. Sibilant geminates are found in the juncture of nominal *-s* stems plus the locative plural *-su* ending, as in *haviṣ-ṣu*, or *nis-* preverb plus *s-* stems, as in *niṣ-ṣidh-* for example (Kobayashi 2004: 45).

The most frequent historical origin of geminates on a cross-linguistic basis is through total assimilation between consonants, and many geminates are formed in this way in Gāndhārī (§ 8.3.1.1 for example) and other MIA languages. This is seen for example in Pali *sakka-* (PTSD s.v. adj. “able, possible”), which is formed through diachronic assimilation of the *-ky-* cluster in Sanskrit *śakya-*, where the geminate *-kk-* contrasts with short *-k-* in Pali *saka-* (adj. “own”). This example points towards the development of homographs in the Kharoṣṭhī orthography, which does not record geminate consonants as such (§ 5.1). Alternately, if developed geminates are non-contrastive there may be little diachronic imperative to retain gemination in the spoken language, beyond any prescriptive grammatical or metrical constraints.

Another interesting source of apparent gemination in Sanskrit is through ‘doubling’ rules which are described in the ancient grammatical treatises, and prescribe or

¹¹³ Refer to Table A3.1 in Appendix A3.

allow the doubling of post-vocalic consonants in many consonant clusters, as in TC > TTC for example¹¹⁴. While the motivations for such doubling are obscure, it may in some cases represent a record of the phonetic lengthening of cluster consonants in particular situation, as for example across a syllable boundary in order to strengthen the coda of the prior syllable, which may also occur for prosodic reasons (Kobayashi 2004: 32-34). While such doubling has been denied by some grammarians, and rules exist which cancel the doubling in certain circumstances, it has been recorded in many traditional manuscripts and inscriptions. It is however normally omitted in contemporary grammars, with the exception of the doubling of *-ch-* > *-cch-* after vowels.

Geminated stops can vary in length between one and a half and three times the length of short stops in careful speech and are mostly limited to intervocalic word medial positions, while word-initial long-stops appear to be rare (Ladefoged and Maddieson 1996: 92). While consonants having all manners of articulation are found in geminates, cross-linguistic surveys suggest that geminates of consonants having lower sonority are more disposed to develop length contrasts than more sonorant sounds. While exceptions exist, this manifests in the order of voiceless stops being most prone to develop length contrasts, followed by voiceless fricatives, nasals, liquids and then glides as the least likely, while voiced stops are less likely than voiceless stops to form contrastive geminates (Gordon 2016: 55, Figure 3.11). It is suggested that perceptual factors play a role in these tendencies in that length contrasts are less perceptible when consonants are acoustically more like surrounding vowels, with glides least likely to participate in length contrasts (Gordon 2016: 56-57). It is also harder to maintain voicing over the extended duration of a voiced geminate stop, where increased articulatory effort is required to overcome aerodynamic factors (Gordon 2016: 154), in contrast with pairs of non-homorganic voiced stops where the movement of the articulators appears to allow the voicing to emerge. Word medial geminates usually form as coda-onset pairs, while also shortening the preceding vowel to some extent (Ladefoged and Maddieson 1996: 92). At the same time, geminates are typically produced with a single but lengthened articulatory movement (pp. 92-93), suggesting that the perception of a coda-onset syllable structure relies on the timing of the lengthened closure rather than a double movement of the articulators.

¹¹⁴ E.g., Whitney 1924: 79 [§229]; Cho 1999: 152-156; Kobayashi 2004: 31-34; Suzuki 2015.

5. The Kharoṣṭhī script and Gāndhārī phonological system¹¹⁵

Similar to the Devanāgarī alphabet used in the writing of Sanskrit, the Kharoṣṭhī script uses an abugida system, with each sign representing a syllabic unit or *akṣara* comprising a consonant plus the neutral vowel *a*, with other vowels indicated by diacritic marks. Unlike other indigenous Indian writing systems, it is however written from right to left, consistent with its Aramaic origins. While the script incorporates a number of punctuation signs, these appear to be used variably by different scribes (Glass 2000: 143-144), and words are typically not separated by spaces within the text, which reads continuously between any punctuation marks. Similar to the Indian Brāhmī script, Kharoṣṭhī was developed for the recording of a MIA language, and does not include all of the characters necessary for the recording of OIA languages¹¹⁶. While there is a broad correspondence between the Brāhmī and Kharoṣṭhī alphabets, the Kharoṣṭhī used in Gandhāra is more limited in not having signs or diacritics to record long vowels and diphthongs, nor to indicate geminate consonants¹¹⁷. The limited marking of phonetic features in the script means that many *akṣaras* have the potential to express a number of underlying pronunciations, as in *-pa-* which might represent any of *-pa-*, *-pā-*, *-ppa-* or *-ppā-* for example.

The sign inventory of the Kharoṣṭhī is known through its own *a-ra-pa-ca-na* alphabet sequence, named after the initial five letters of the alphabet¹¹⁸. Even so, the Kharoṣṭhī graphemes represent an attempt to capture the sounds of a MIA language which would have had many similarities to those in Sanskrit, and for ease of comparison the Gāndhārī phonological system can also be arranged on a similar basis to the Sanskrit *varṇamālā*. While correlations to the Sanskrit pronunciation (§ 4.1) are available through parallel texts and word forms, no historic grammatical or phonetic descriptions are known for Gāndhārī, similar to those available for Sanskrit, and which might guide a better understanding of the actual pronunciation of the characters. Along with the absence of ongoing spoken or recited tradition, this means that sounds may only be assumed within

¹¹⁵ This section draws extensively from the comprehensive account of the Kharoṣṭhī script prepared by Glass (2000), supplemented by later publications which describe variations in the orthography observed in documents which became available subsequent to that date. Some of the transliteration conventions described by Glass have also been amended in later general use by Gāndhārī scholars.

¹¹⁶ Salomon 2003b: 89; Strauch 2012: 134-136.

¹¹⁷ Although not occurring as independent vowel signs, long vowels are marked on consonant *akṣaras* in the DhP-G^K and ND from Central Asia, and in later stage Bamiyan manuscripts (Glass 2000: 36, 41, 43, 137).

¹¹⁸ Salomon 2003: 92; Strauch 2012: 137; Salomon 2018: 90.

the broader phonetic classifications established by Sanskrit parallels. The origin of the language from an earlier Ṛgvedic dialect (§ 2.1), combined with regional patterns of dialect development, further complicate any attempts at more precise classification of the sounds. Even so, the use of many diacritic modifiers in the Kharoṣṭhī script (§ 5.2) suggest that the scribes were attempting to capture variations or transitions in pronunciation from OIA. These may also reflect Persian and other regional language influences due to the strong cultural interchange that occurred in the region.

The Gāndhārī phonological system is presented in Table 5.0, with the consonants organised in a similar manner to Table 4.1 for the Sanskrit system, following the IPA format. The table includes the assumed IAST transliteration and a representative rendering of the Kharoṣṭhī script for each *akṣara*. This is based on the hand of British Library Scribe 1 (BLS1), as collated by Salomon (1999: 111, Fig. 12)¹¹⁹, which provides a reasonably clear representation of the basic characters. The orthographic styles and practices used in the writing of characters by different scribes are quite variable, however, and renderings of the script can differ markedly from the forms in the table. This is seen, for example, in the charts of the script as written by various scribes presented by Glass (2000: 161-187, Appendix B). These variabilities appear to depend on a number of factors such as the speed of writing, the quantity of ink in the pen, natural variabilities in the base material, and the purpose of the document. The written forms of *akṣaras* having similar shapes, such as *ta* and *da*, *śa* and *ya* or *va* and *a* for example, can also bear a close resemblance as a result of these variabilities, and textual experience and contextual interpretation are often required to discern the intended form. And while the cursive development and forms of some of the Kharoṣṭhī characters did change over time, the basic forms of most of the *akṣaras* were quite stable over the six centuries or so of known use, with little significant variation occurring between the older and later forms (Salomon 2003: 94).

Various features of the Kharoṣṭhī script are described in more detail in the following sections. These are not intended to be comprehensive in the sense of describing all of the variations that have been observed, but rather to provide a broad descriptive representation some of the main variations which relate to linguistic understanding of the language.

¹¹⁹ The symbols for *ga* and *na* are not recorded in the work of BLS1, and these characters are taken from the hands of BLS9 and BLS21 respectively (Glass 2000: 167, 173).

Table 5.0: The Gāndhārī Phonological system showing a representative version of the Kharoṣṭhī script

Consonants	Labial		Dental		Retroflex		Palatal		Velar		Glottal
	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	voiced
Stops											
unaspirated	<i>pa</i> 𑀧	<i>ba</i> 𑀢	<i>ta</i> 𑀭	<i>da</i> 𑀮	<i>ṭa</i> 𑀓	<i>ḍa</i> 𑀕	<i>ca</i> 𑀇	<i>ja</i> 𑀉	<i>ka</i> 𑀫	<i>ga</i> ⁽¹⁾ 𑀔	
aspirated	<i>pha</i> 𑀧𑀖	<i>bha</i> 𑀢𑀖	<i>tha</i> 𑀭𑀖	<i>dha</i> 𑀮𑀖	<i>ṭha</i> 𑀓𑀖	<i>ḍha</i> 𑀕𑀖	<i>cha</i> 𑀇𑀖	(<i>jha</i>) 𑀉𑀖	<i>kha</i> 𑀫𑀖	<i>gha</i> 𑀔𑀖	
Nasals ⁽¹⁾		<i>ma</i> 𑀢𑀭		<i>na</i> ⁽¹⁾ 𑀮𑀭		<i>ṇa</i> 𑀕𑀭		<i>ña</i> 𑀉𑀭		(<i>ṇa</i>) 𑀕𑀭	
Fricatives			<i>sa</i> 𑀭𑀖		<i>ṣa</i> 𑀕𑀖		<i>śa</i> 𑀇𑀖				<i>ha</i> 𑀬
Semi-vowels tap approximants		<i>va</i> 𑀢𑀭		<i>la</i> 𑀮𑀭		<i>ra</i> 𑀕𑀭		<i>ya</i> 𑀉𑀭			
Initial vowels:	<i>ā</i> 𑀢	𑀢	<i>ī</i> 𑀮	𑀮	<i>ū</i> 𑀕	𑀕	<i>e</i> 𑀇	𑀇	<i>o</i> 𑀉	𑀉	
Diacritic vowels:	<i>ka</i> 𑀫	𑀫	<i>ki</i> 𑀫𑀖	𑀫𑀖	<i>ku</i> 𑀫𑀭	𑀫𑀭	<i>ke</i> 𑀫𑀭	𑀫𑀭	<i>ko</i> 𑀫𑀭	𑀫𑀭	

Source: Symbols copied from Salomon (1999: 111, Figure 12) [⁽¹⁾ *ga* and *na* copied from Glass 2000 (167, 173 Tables A3 and A4)]

5.1 The use of syllabic modifiers in the Kharoṣṭhī Script


Diacritic marks added to the basic characters are a common feature of the Kharoṣṭhī script. These act as syllabic modifiers, and appear to have been used in three main contexts (Glass 2000: 119, 133-136):


- To indicate phonetic modifications of intervocalic consonants.
- To indicate the modified pronunciation of *akṣaras* which are reflexes of OIA consonant clusters.
- To indicate some of the retained OIA consonant clusters.









Furthermore, the characters of the Kharoṣṭhī script are often written with flourishes to the bottom of the vertical stem (Glass 2000: 21). While the appearance of these ‘foot marks’ can be similar to other diacritic marks, the patterns of their occurrence suggest they are most likely scribal flourishes rather than markers of phonetic modification or consonant clusters.

5.2 Phonetic modification of intervocalic consonants

Diacritics which indicate the phonetic modification of single intervocalic consonants are described in this section. These are distinct from other markings which appear to indicate the reflexes of OIA consonant clusters (§5.4). The modified pronunciation of intervocalic consonants is mostly indicated with a rightward projection from the base of their stem, resembling a tail, or ‘cauda’. In most cases this diacritic has been interpreted to indicate a weakened or fricativised pronunciation (Glass 2000: 136), which is discussed further in § 8.2. The modification is conventionally transliterated by the addition of an underscore mark to the IAST character, although an acute accent diacritic above the character was also used in an earlier convention. These modified forms often occur as one of several alternate spellings written for a particular reflex, even by the same scribe (§ 7.2). The most common forms recorded in the Gandhāran source materials include:

- *ka*: This modification has been used by the RS scribe to represent some reflexes of Sanskrit *ka* in intervocalic situations, seen in the form  in the RS-5, RS-14 and

RS-20 manuscripts¹²⁰. It does not appear to be used extensively however, and nor is it well attested elsewhere. This character is similar to the conjunct consonant character for *kra* , written here in the hand of BL scribe 1 (§ 5.4.2.1). This form does not appear to have been used to mark reflexes of intervocalic *ka* in the RS 12 manuscript however, where it is otherwise used in the marking of the *kra* conjunct (Silverlock 2015: 190-191).

- *ga*: This modification has been used in a wide range of documents to represent reflexes of Sanskrit *ga* in intervocalic situations. These include extensive use in the RS-5, RS-12 RS-20 and BC manuscripts¹²¹, seen in the forms  written by the RS scribe and  by the BC scribe 5¹²², and although less extensively, in many other source materials. These include the InsK, the ND, the KDhp as well as some of the MS and BL manuscripts, seen in the form  written by the BL scribe 9 for example¹²³. These forms are similar to the conjunct character for *gra* (§ 5.4.2.1), which are written as ,  and  by the RS scribe, BC scribe 5 and BL scribe 9 respectively¹²⁴. While the distinction between a more curved or straight base stroke in the forms of *ga* and *gra* is generally clear in the hand of individual scribes, these strokes are not always rendered consistently between scribes, and in some cases similar forms appear for both for *ga* and *gra*¹²⁵.
- *ḍa*: This modification has been used extensively in the RS manuscripts to represent reflexes of Sanskrit *ta* and *da* in intervocalic situations, seen in the typical form  written by the RS scribe in the RS-5, RS-12, RS-14 and RS-20 manuscripts for example¹²⁶. Single uses have also been seen in the Theodoros casket in the InsK as well as by the BL scribe 21, written as  in the latter case¹²⁷. While it is not clear whether this represents a phonological modifier or scribal flourish in the hand of this scribe, intervocalic *ḍa* appears to be distinguished from *da* in the RS collection, suggesting lenition of the single consonant in intervocalic positions. This form is also

¹²⁰ Glass 2007: 92 [Table 14], 95, 114; Salomon 2008: 338; Marino 2017: 116-118.

¹²¹ Glass 2007: 114-115; Silverlock 2015: 192-193; Marino 2017: 116-119; Schlosser 2022: 81-82.

¹²² Glass 2007: 92, Table 14; Schlosser 2022: 62, Table 7.


¹²³ Glass 2000: 58, 167, Table 3.





¹²⁴ Glass 2000: 171, Table 3; Glass 2007: 93, Table 14; Schlosser 2022: 62, Table 7.

¹²⁵ Contextual interpretation of the palaeography is often required due to such variabilities in the written forms.





¹²⁶ Glass 2007: 92 [Table 14], 116; Salomon 2008: 339; Silverlock 2015: 193-195; Marino 2017: 121-122.

¹²⁷ Glass 2000: 79-80, 173, Table 4.

similar to the conjunct character *dra*  (§ 5.4.2.1), and contextual interpretation is necessary to distinguish between these two similar forms.

- *śa*: This apparent modification of the dental sibilant *sa* has been used in a wide range of documents to represent reflexes of the Sanskrit aspirated dental stops *tha* and *dha* in intervocalic situations, as well as in reflexes of the dental sibilant itself. These include uses by the RS scribe in manuscripts such as RS-5, RS-12, RS-14 and RS-20 for example, as well as uses by the BC scribe and some of the BL scribes¹²⁸. Sometimes referred to as “curly s”, this modification is seen for example in the forms  written by the RS scribe,  written by the BC scribe,  written by BL scribe 9, and  written by BL scribe 21¹²⁹. This form is also seen in the Bajaur casket and Wardak vase inscriptions in the InsK, as well as in the ND. While the phonetic meaning is not universally agreed, one suggestion is that it may represent a voiced fricative pronunciation, which would develop through lenition of the aspirated stops and voicing of the original unvoiced sibilant¹³⁰. It is also seen in alternation with the Sanskrit genitive singular ending *-śya*.

The following forms of this type are also recorded in more limited situations:

- *ja*: This modification appears to represent a fricative form of *ja* regularly corresponding to Sanskrit *śa*, as in Skt. *avakāśa* > G *avagaśa*. This form has been found in the ND as well as the InsK in the Bajaur casket, seen for example in the form  (Glass 2000: 66).
- *ta*: This modification is seen in some of the InsK, as for example in *tī*  from the Bajaur casket. It does not appear to occur in any of the Gandhāran manuscripts however, and could possibly reflect lenition to a dental fricative (Glass 2000: 77)
- *śa*: This modification does not appear widely, and is seen in the form  in the Bajaur casket in the InsK, and  in the BC manuscripts by the BC scribe 5¹³¹. In the latter

¹²⁸ Glass 2000: 107-109; Glass 2007: 114, 116-117; Salomon 2008: 339; Silverlock 2015: 197-200; Marino 2017: 121-122, 126; Schlosser 2022: 84-87.

¹²⁹ Glass 2000: 170 [Table 3], 174 [Table 4]; Glass 2007: 93, Table 14; Schlosser 2022: 65, Table 7.

¹³⁰ Glass 2000: 108-109.

¹³¹ Glass 2000: 100; Schlosser 2022: 65, Table 7.

case it appears in reflexes of the Sanskrit palatal sibilant $\acute{s}a$ ¹³². This appears to indicate a phonetic difference such as a voicing of the sibilant for example¹³³.

- $\underline{s}a$: This form does not appear widely, seen as 𑖦 in the BC manuscripts by the BC scribe 5, and also reported in the RS-12 manuscript. It is used to record reflexes of the Sanskrit retroflex sibilant \acute{s} , and appears to indicate voicing of the sibilant¹³⁴.

A small number of other diacritics have only been observed in the KDhp and ND, possibly reflecting some unique Central Asian forms. These include:

- $\bar{g}a$: This modification combines a horizontal bar above the character along with a rightward underscore, as for example in the form 𑖧 from the KDhp, where it possibly represents aspiration, or 𑖨 from the ND where it appears to represent a following nasal (Glass 2000: 59).
- γa : This character is only found in ND in the form 𑖩 , and may represent the bi-labial approximant [w], although its use is inconsistent (Glass 2000: 98).
- Virāma: Although Sanskrit word-final consonants are typically lost in Gāndhārī (§ 8.2.1.3), final consonants without vowels have been observed in the ND (Glass 2000: 138). These are represented with a subscript in the Kharoṣṭhī script, which appears to be the equivalent of the Devanāgarī virāma.
- Infinitas: A sign resembling the infinity symbol written beneath some characters in a few of the ND, which may represent a vowel modifier of some kind (Glass 2000: 138).

5.3 Consonant clusters in the Kharoṣṭhī script

Many of the OIA consonant clusters are also recorded in Gāndhārī. Recorded clusters may not always represent the most advanced form of the language, however, and the reflexes of OIA clusters may be reproduced in conservative forms for various reasons in some source documents (§ 8.3). Hence the recording of a particular cluster in Kharoṣṭhī does not necessarily confirm its widespread use or general retention in the language. Consonant clusters are indicated in four ways in the Kharoṣṭhī script:

¹³² Schlosser 2022: 81, 84.

¹³³ As discussed by Baums (2007: 150).



¹³⁴ Silverlock 2015: 229, 244-245; Schlosser 2022: 65, Table 7.

- Through the use of additional diacritic marks to the *akṣaras* for individual consonants.
- By the use of *akṣaras* which appear to have evolved their own unique forms.
- By the use of *akṣaras* which are formed by the evident combination of the characters for the two conjunct consonants to form special ligatures.
- Clusters are also interpreted in some contexts. This occurs primarily when intervocalic stops are recorded in a non-reduced form, in contrast with reflexes which indicate the lenition of single intervocalic consonants as fricatives or approximants (§§ 5.2, 8.2). Recording of the strong form may indicate the retention of an original or developed stop-stop geminate (§ 8.3.1) or nasal-stop cluster (§ 8.3.2). Lenition of intervocalic consonants is not always indicated however, and it is often necessary to review the intention by comparison with parallel Sanskrit forms.







As noted in § 5.2, certain diacritic marks are also used to indicate the modified pronunciation of *akṣaras* which are reflexes of OIA consonant clusters. These may reflect partial traces of the original cluster and are subsequently dealt with in this section.

5.3.1 Characters which record reflexes of original OIA clusters

Several characters appear to represent consonants which are reflexes of an underlying consonant cluster, and reflect or retain a modified form of its pronunciation in some way (Glass 2000: 136). These are marked in two ways - with either a superscript horizontal bar or an upward diagonal acute accent diacritic, both of which are reflected in the IAST transliteration. The phonetic interpretation of these *akṣaras* is discussed further in § 8.3. The recorded forms marked with a superscript horizontal bar include:

- *ṁa*: This character is recorded in the BC manuscripts as well as the later Schøyen collection and the ND, as for example in the form  by the scribe 2 (Glass 2000: 62, 185 [Table 7]). This form regularly corresponds with Skt. *-śca*, as in Skt. *paścāt-* > G *paṁa-* for example, where it appears to represent a geminate reflex (Schlosser 2022: 90 [Table 10], 97), although this modification is not included in *-c-* reflexes seen in other materials (§ 8.3.3.1).
- *ṁa*: This character is seen in the BL manuscripts and Schøyen collection as well as the Dhp-G^K, as for example in  in the hand of BL scribe 9. It also appears to have

occurred in embryonic form as a dot in the Aśokan inscriptions¹³⁵. This character appears to correspond to Sanskrit *dhy-* as in Skt. *dhyānam* > G *ḥaṇo* and *-madhye* > *-maḥe* for example, where it appears to represent an aspirate or geminate stop (Salomon 2000: 88).

- *mā*: This character, written as , has been found only three times on fragment 5A of the BL manuscripts, where it corresponds to Skt. *-hm-*¹³⁶. This is normally resolved as the *-(m)m-* geminate (§ 8.3.3.2), suggesting that the diacritic is intended to represent the geminate nasal.
- *ṣā*: This character has been found in some of the InsK and in the BL and BC manuscripts, as for example in the form  written by BL scribe 9 and  by BC scribe 5. It is used in locations corresponding with Skt. *ṣṇ-*, as in Skt. *kṛṣṇam* > G *kriṣo* and *ūṣṇa* > *uṣa* for example¹³⁷. Although the phonetic meaning is uncertain, this has been interpreted to represent retention of the cluster.
- *śā*: This character has been found in some of the BL manuscripts, and also in one of the Schøyen fragments as well as the ND, seen for example in the form  from the Schøyen collection. It occurs in syllables corresponding to Skt. *śn-*, as in Skt. *śneha-* > G *śeḥa-*, and also for *sn-*, as in Skt. *snāyati-* > G *śayadi-* for example¹³⁸. Although the phonetic meaning is uncertain, this has been interpreted to represent retention of the cluster.
- The addition of a superscript to the palatal sibilant *śa*, as in *ṣ́a*, has been observed in the BL and SC manuscripts, along with one fragment from the Pelliot collection, seen for example in the form  recorded on BL fragment 10. The superscript is thought to have been used to distinguish the character from the similar form of *ya*, and is considered to be a graphic device only¹³⁹. A variant form of *śa* having a curled footmark on the base of either leg, as in  by BL scribe 9 for example, can be

¹³⁵ Glass 2000: 65; Salomon 2000: 58, Table 2.

¹³⁶ Glass 2000: 93; Salomon 2000: 220. Although recorded as part of frame 15 of BL8 by Glass, the fragment on which this character is recorded is referred to as fragment 5A by Salomon. Refer also to Salomon (1999: 45).



¹³⁷ Glass 2000: 103; Salomon 2000: 61 [Table 2], 91; Lenz 2003: 131; Baums 2007: 173; Schlosser 2022: 65 [Table 7], 96 [Table 11].

¹³⁸ Glass 2000: 109; Salomon 2000: 78; Baums 2007: 173.




¹³⁹ Glass 2000: 99; Falk and Karashima 2015: 57-58.

transliterated as $\acute{s}a$, and when combined with a superscript may also rendered as $\acute{s}a^{\bar{}}$, both of which have the same phonetic intention as $\acute{s}a$ ¹⁴⁰.

The recorded forms marked with an upward diagonal acute accent diacritic include:

- $\acute{k}a$: This character is recorded in the ND and PC from Central Asia, in the InsK on the Kurram casket, in the Schøyen collection, and in the BL manuscripts in the hand of BL scribe 1. It is seen for example in the form  from the Pelliot collection. This often corresponds to Sanskrit *sk*, as in Skt. *saṃskāra-* > G *saṃkara-* and *puraskṛtaḥ* > *purakidu* for example, where it appears to represent a geminate reflex¹⁴¹.
- $\acute{t}ha$: This modification is found in the KDhp, the ND, the Aśokan inscriptions, and the BC, BL and RS manuscripts. It is marked by an upward extension to the left-hand arm of the character, seen for example in  written by BL scribe 1. In the manuscript source documents it has been used to record reflexes of *-ṣṭh-*, *sth-* and *-sth-*, as in BHS *upatiṣṭhatha-* > G *vaṭhasa-*, Skt. *sthita-* > G *ṭhida-* and *upasthita-* > *upaṭhita-* for example¹⁴². There appears to be a complex phonetic relationship between this character and $\acute{t}ha$, *sta* and *tha*¹⁴³, which is discussed further in § 8.3.3.1.

The following cases involving an overscore diacritic are only recorded in the Central Asian source materials:

- $\bar{g}a$: This character is seen in both the KDhp and ND. In the latter it possibly represents a following *n* in Sanskrit, as in Skt. *nagna-* > G *naḡa* (Glass 2000: 59).
- $\bar{n}a$: This character is only found in the KDhp, seen in the form , where it corresponds to OIA *ndh* (Glass 2000: 82).
- $\bar{h}a$: This character has only been observed once in the ND, seen in the form  for $\bar{h}e$, and appears to correspond to OIA *hṇ* (Glass 2000: 113).
- $k\bar{s}a$: This modification of the character for the conjunct consonant *kṣa* is seen in the ND in the form , and possibly indicates a development of the *kh* or *ch* reflexes (Glass 2000: 116).

¹⁴⁰ Glass 2000: 168, Table 3; Falk and Steinbrückner 2020: 8.

¹⁴¹ Salomon 1998a: 156, Fig. 1; Glass 2000: 52-53; Salomon 2008: 126.

¹⁴² Glass 2000: 71-72; Allon 2001: 95; Salomon 2008: 94 [Table 7], 125-126; Schlosser 2022: 96, Table 11.













¹⁴³ E.g., Brough 1962: 75-77; Baums 2009: 164-166; Silverlock 2015: 270-271.

5.3.2 Conjunct consonants recorded using diacritic marks



Consonant clusters involving semi-vowels are typically indicated with diacritic marks in the Kharoṣṭhī script, including post-consonantal $C\{v, y, r\}$ as well as pre-consonantal rC clusters (Glass 2000: 119-127). Diacritics are also used for the marking of post-consonantal $C\{p, y\}$, but this is only seen in the ND. The diacritics take various forms which typically appear as additional foot strokes, and must be distinguished contextually from syllabic modifiers (§ 5.3) and other non-phonemic scribal flourishes having similar forms.

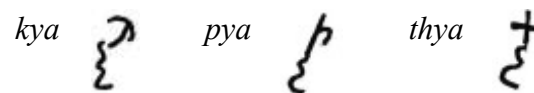
5.3.2.1 $C\{v, y, r\}$ clusters





Post-consonantal v : This is typically marked with an upward hook rising to the right from the base of the radical stem. This marking appears to have been used consistently during all periods and in source materials from both Central Asia and Gandhāra (Glass 2000: 125-126). The following table shows some of the forms of this character seen in the Gandhāran manuscript sources:

<i>tva</i>	 BLS1 (Allon 2001: 59)	<i>kva</i>	 BLS1 (Salomon 2008: 94)
	 BLS2 (Lenz 2010:18)	<i>dva</i>	 BLS9 (Salomon 2000: 62)
	 BLS9 (Salomon 2000: 62)		 RS-5 (Glass 2007: 93)
	 RS-5 (Glass 2007: 93)	<i>sva</i>	 BCS5 (Schlosser 2022: 65)
	 BCS5 (Schlosser 2022: 63)		
<i>tvo</i>	 BLS1 (Salomon 2008: 94)		
<i>tvi</i>	 BCS5 (Schlosser 2022: 63)		
<i>tve</i>	 BCS5 (Schlosser 2022: 63)		

The *tva* cluster appears to occur most frequently, and is also seen with other vowel diacritics, while *kva*, *dva* and *sva* appear to occur less frequently. The upward stroke of the hook is typically quite strong and distinct from other flourishes and diacritics attached to the base of the radical stem, although a shorter stroke has been observed in some manuscripts. A number of *rva* clusters also occur, but these are marked in a similar way to other pre-consonantal rC clusters and are dealt with in § 5.4.2.2 below.

Post-consonantal *y*: While OIA C + *y* clusters are mostly assimilated in Gāndhārī reflexes (§§ 8.3.1.4.3, 8.3.2.4.3, 8.3.3.4.3), post-consonantal *y* has been retained in some instances, mostly in the inscriptional sources as well as some of the Bamiyan fragments. Post-consonantal *y* is represented by various diacritic marks attached to the base of the consonantal akṣara. These include the writing of a normal *y* sign beneath the prior consonant, such as in  for *mya* from the third century BCE Aśokan inscription at Mānsehra, and more commonly by the attachment of a wavy line or a hook to the base of the radical stem, which is also detached in some cases (Glass 2000: 119-121). Examples of a wavy line include  for *bhye* from the Aśokan inscriptions, and several examples from the second century CE Wardak vase shown below¹⁴⁴.























Examples of a hooked stroke include  for *tya* where the hook appears to be detached, and  for *vya* where the hook is attached, both from the third century CE Bamiyan manuscripts (Glass 2000: 119-121). A character which has been contextually interpreted as representing *vya* and *vyo* (§ 8.3.4.2.4) has been observed in preliminary studies of the BL fragment 4 by Tim Lenz. The scribe has been identified as ‘slanted hand’ based on his writing style. The characters, which had not been previously observed in the Gandhāran avadānas are rendered as  *vya* and  *vyo*¹⁴⁵, involving a more continuous single hook diacritic (T. Lenz, personal communication, September 11 and 18, 2023). Another example of the apparent retention of *vy-* involves the writing of *vi-* using a simpler vowel diacritic marker across the akṣara, in an apparent graphic shorthand which relies on the phonetic similarity between these sounds (§ 8.3.4.2.4). A number of *rya* type clusters also occur, but these are marked in a similar way to other pre-consonantal *rC* clusters and are dealt with in § 5.4.2.2 below.


Post-consonantal *r*: This is typically marked with a stroke to the right from the base of the character, although the length and shaping of the stroke varies between source materials. This marking appears to have been used consistently during all periods and in both the Central Asian and Gandhāran source materials (Glass 2000: 124-125). The only

¹⁴⁴ These have been traced from the image of the Wardak vase in Konow (1929: Plate XXXIII, facing p.170).



¹⁴⁵ These versions have been traced from images of the manuscript forwarded by Tim Lenz.

exception is the marking of *mra* with a curved stroke added to the upper right stem of the character. The following table shows some of the forms of this character seen in the Gandhāran manuscript sources:

<i>kra</i>	 BLS1 (Allon 2001: 59)	<i>pra</i>	 BCS5 (Schlosser 2022: 64)
<i>gra</i>	 BCS5 (Schlosser 2022: 62)		 BLS2 (Lenz 2003: 117)
	 BLS9 (Salomon 2000: 62)	<i>bra</i>	 BLS1 (Allon 2001: 59) b
<i>ghra</i>	 BLS1 (Allon 2001: 59)	<i>mra</i>	 BCS5 (Schlosser 2022: 64)
<i>tra</i>	 BCS5 (Schlosser 2022: 63)		 BLS9 (Salomon 2000: 62)
	 BLS1 (Allon 2001: 59)		 BLS2 (Lenz 2003: 117)
<i>dra</i>	 BLS9 (Salomon 2000: 62)	<i>vri</i>	 RS-5 (Glass 2007: 93)
<i>dru</i>	 BCS5 (Schlosser 2022: 63)	<i>vru</i>	 BLS2 (Lenz 2003: 117)
<i>dhra</i>	 BLS1 (Allon 2001: 59)	<i>śra</i>	 BLS2 (Lenz 2010: 18)
<i>dhro</i>	 BLS1 (Allon 2001: 59)	<i>sre</i>	 BLS1 (Salomon 2008: 95)

The similarities between these diacritics and those used to mark phonetic modifications to single consonants (§5.3) are evident. The main distinguishing point appears to be that these diacritics are drawn in a continuous curve integral with the base of the stem in some cases, although others spring from an angular junction having more similarity with those described in § 5.3, and distinguishing between these diacritics requires contextual interpretation. Another similar character is seen the writing of *mi* as  in the Wardak vase in the InsK, with an underscore attached to the base of the vowel diacritic. This suggests interpretation as post consonantal *r*, and has been transcribed as *m(r)i* by Konow (1929: 170).

5.3.2.2 *r*-C clusters

Pre-consonantal *r* is typically marked with a clockwise spiral loop formed from the base of the radical stem in the manuscript sources, written in either a partial or near to complete form in different documents. An earlier version involving a horizontal stroke through the middle of the stem, as for example in the forms  or  for *rva*, is seen in the earlier Aśokan inscriptions (Glass 2000: 121-124). The following table shows some of the forms of this character seen in the Gandhāran source materials. Variants on the spiral loop

seen in the BL manuscripts include *rga* ੜ compared with *ga* ੜ, *rśa* ੜ compared with, *śa* ੜ, and *rya* ੜ compared with *ya* ੜ for example, and *rma* ੜ compared with *ma* ੜ in which the loop is added to the top of the character (Glass 2000: 162-171 [Tables 2 and 3]). The clockwise loop distinguishes this form of character from the post-consonantal semi vowel clusters discussed in the previous section:

<i>rkha</i> ੜ BLS1 (Salomon 2008: 95)	<i>rma</i> ੜ BCS5 (Schlosser 2022: 64)
<i>rga</i> ੜ BCS5 (Schlosser 2022: 62)	<i>rmo</i> ੜ BCS5 (Schlosser 2022: 64)
<i>rga</i> ੜ BLS9 (Salomon 2000: 62)	<i>rma</i> ੜ BLS1 (Allon 2001: 59)
<i>rgha</i> ੜ BCS5 (Schlosser 2022: 62)	<i>rmu</i> ੜ BLS1 (Allon 2001: 59)
<i>rce</i> ੜ BLS1 (Salomon 2008: 95)	<i>rma</i> ੜ BLS2 (Lenz 2003: 117)
<i>rji</i> ੜ BCS5 (Schlosser 2022: 62)	<i>rva</i> ੜ BLS2 (Lenz 2003: 117)
<i>rṭa</i> ੜ Shāhbāzgarhī (Glass 2000: 123)	<i>rva</i> ੜ BLS1 (Salomon 2008: 94)
<i>rti</i> ੜ BLS9 (Salomon 2000: 62)	<i>rva</i> ੜ RS-5 (Glass 2007: 93)
<i>rtha</i> ੜ BCS5 (Schlosser 2022: 63)	<i>rva</i> ੜ BCS5 (Schlosser 2022: 64)
<i>rdha</i> ੜ BLS9 (Salomon 2000: 62)	<i>rva</i> ੜ Mānsehra (Glass 2000: 124)
<i>rpa</i> ੜ BCS5 (Schlosser 2022: 64)	<i>rvo</i> ੜ BLS1 (Allon 2001: 59)
<i>rbha</i> ੜ BLS1 (Allon 2001: 59)	<i>rya</i> ੜ BLS9 (Salomon 2000: 62)
<i>rno</i> ੜ BLS2 (Lenz 2003: 117)	<i>rya</i> ੜ BLS2 (Lenz 2010: 18)
<i>rṇa</i> ੜ BLS2 (Lenz 2010: 18)	<i>rya</i> ੜ BCS5 (Schlosser 2022: 64)
<i>rṇo</i> ੜ BLS1 (Allon 2001: 59)	<i>rśa</i> ੜ BLS1 (Allon 2001: 59)
	<i>rśi</i> ੜ BLS1 (Salomon 2008: 95)
	<i>rṣa</i> ੜ BLS1 (Salomon 2008: 95)

5.3.2.3 Anusvāra

The use of *anusvāra* in the Kharoṣṭhī script is confirmed by the marking of consonants in locations where *anusvāra* is anticipated through OIA parallels, although this is not applied consistently in all of the source materials. It is typically marked with an additional foot stroke, similar to the *ma* character, added to the base of characters (Glass

2000: 134-135). Although this suggests a form of a combined character, similar to the Sanskrit marking it indicates a post-vocalic nasalisation at the end of the syllable, rather than a conjunct consonant within the marked syllable, and in this context is marking a form of NC cluster in the following syllable (§ 4.3).



Anusvāra appears to have been used with some regularity in the inscriptional materials, as for example in *aṃ* (𑀅) from the Aśokan inscriptions at Shāhbāzgarhī (Glass 2000: 134-135). Among the Gandhāran manuscripts it is not used regularly by many of the scribes, although it appears to have been used consistently by BL scribe 9, as in *aṃ* 𑀅, *paṃ* 𑀅, *taṃ* 𑀅, *chaṃ* 𑀅, *yaṃ* 𑀅, and *saṃ* 𑀅 for example (Salomon 2000: 61, Table 2), as well as by Schøyen scribe 2, seen in *kaṃ* 𑀅, *taṃ* 𑀅, *yaṃ* 𑀅, and *saṃ* 𑀅 for example (Glass 2000: 187, Table 7). In the BL manuscript examples the footmark is continuous with the main character, while in the Schøyen manuscripts it is detached, demonstrating variations in the writing style between scribes. This marking also occurs in situations where it is not anticipated by context, and has been described as ‘pseudo-*anusvāra*’ in these situations. This usage may be the result of graphic confusion (Salomon 1999: 211; Glass 2000: 134). It is also possible for the *anusvāra* diacritic to be confused with loose scribal flourishes or other diacritics such as those used for the *rC* clusters.











5.3.3 Conjunct consonants recorded using unique characters

The conjunct consonants *kṣa* and *sta* are recorded in the Gandhāran source materials with unique ligatures, which appear to bear no resemblance to the component characters (Glass 2000: 115-119). Some examples of the writing of these characters in the Gāndhārī manuscripts are tabulated below. It is unclear whether these characters developed as modifications to single consonant characters or were ad hoc improvisations (Salomon 1998b: 49). Both examples involve combinations of sibilants and stops, and there is some uncertainty about the phonetic nature of *kṣa*, which was possibly pronounced as an affricate in Sanskrit¹⁴⁶, and the syllabic patterning of *sta*-¹⁴⁷, which may have contributed to their development as unique ligatures in the Kharoṣṭhī script. It is possible to imagine the *kṣa* ligature evolving from the basic form of the *ṣa* character 𑀓, modified

¹⁴⁶ A unique ligature is also used for *kṣa* in the Devanāgarī script.

¹⁴⁷ Refer to the discussion about Sanskrit syllables in Appendix A4.



through rotating the ‘horn’ shaped head upward. While the *mra*  cluster has been classified with the other unique ligatures by Glass (2000: 117), it is recorded using a diacritic attached to the upper right corner of the cup shaped *ma* character, which does not lend itself to the more common stroke springing from the base of the stem used in other *Cr* clusters (§ 5.4.2.1). Another atypical ligature is used to represent *vha* , but is only recorded in some early coin inscriptions as well as the Central Asian source materials (Glass 2000: 118).





























<i>kṣa</i>  BLS1 (Allon 2001: 59)	<i>sta</i>  BLS1 (Allon 2001: 59)
 BLS2 (Lenz 2003: 117)	 BLS2 (Lenz 2003: 117)
 BCS5 (Schlosser 2022: 62)	 RS-5 (Glass 2007: 93)
 RS-5 (Glass 2007: 93)	<i>sti</i>  BLS21 (Glass 2000: 174)
	<i>stri</i>  BLS1 (Salomon 2008: 95)
	<i>staṃ</i>  BLS9 (Salomon 2000: 62)

5.3.4 Conjunct consonants recorded using combination characters

A significant number of conjunct consonant characters are formed through the evident combination of individual characters (Glass 2000: 127-133). Many of these are only found in the central Asian documents however, and a listing of combination characters used in the Gandhāran source documents is presented in the following table. A number of these characters, including, *kta*, *pta*, *ṣma* and *hme*, which are highlighted in bold in the table, are only found in the later stage early third century CE Schøyen manuscripts¹⁴⁸, and are reflexes of Sanskrit clusters which are normally assimilated or reduced in the reflexes recorded in earlier Gandhāran manuscripts (§§ 8.3.1.1, 8.3.3.2). These appear to represent the restoration of historic spellings in a later stage ‘Sanskritisation’ of Gāndhārī (§ 8.7), suggesting that they were developed by the scribes working in or around Bamiyan around this time or not long before. Other clusters such as *tma* and *ṣki* have also been used in some of the second century CE InsK (§§ 8.3.1.2, 8.3.3.1) as well as the Schøyen manuscripts, but are also absent from the earlier

¹⁴⁸ E.g., Allon and Salomon 2000: 266; Salomon 2001: 244-245; Baums, Glass and Matsuda 2016: 186-187; Baums et al. 2016: 268-269; Jantrasrisalai et al. 2016: 12-13.

Gandhāran manuscripts. Combined characters are typically written with a form of the first consonant character placed above that of the second, and the individual components can be discerned in most cases. However, in some instances, such as *śpa*  written by BL scribe 1 for example, there is a graphic distortion in the writing requiring more careful study and possible contextual interpretation. A character which represents *lpa* (§ 8.3.4.1.1) has been observed in preliminary studies of the BL fragment 4 by Tim Lenz, and is rendered as , involving an obvious combination of the individual characters by the scribe¹⁴⁹, who is identified as ‘slanted hand’ based on his writing style. (T. Lenz, personal communication, September 11 and 18, 2023).

<i>kta</i>  Schøyen 1 (Glass 2000: 184)	<i>śpri</i>  BCS5 (Schlosser 2022: 65)
<i>kṣa</i>  BLS1 (Allon 2001: 59)	<i>ṣki</i>  Schøyen 1 (Glass 2000: 184)
<i>jña</i>  Schøyen 2 (Glass 2000: 187)	<i>ṣtha</i>  Schøyen 2 (Glass 2000: 187)
<i>tta</i>  Schøyen 2 (Glass 2000: 187)	<i>ṣthu</i>  InsK-K (Glass 2000: 133)
<i>tma</i> ¹⁵⁰  InsK-K (Konow 1929: 162-165)	<i>ṣṇa</i>  BLS2 (Lenz 2010: 18)
<i>tsa</i>  BLS9 (Salomon 2000: 62)	<i>ṣma</i>  Schøyen 1 (Glass 2000: 184)
<i>pta</i>  Schøyen 1 (Glass 2000: 184)	<i>sta</i>  BLS1 (Allon 2001: 59)
<i>psya</i>  BLS21 (Glass 2000: 174)	<i>stha</i>  Schøyen 1 (Glass 2000: 184)
<i>mma</i>  Aś-H (Glass 2000: 131)	<i>spa</i>  RS-5 (Glass 2007: 93)
<i>śpa</i>  BLS1 (Allon 2001: 59)	 BCS5 (Schlosser 2022: 65)
 BLS2 (Lenz 2003: 117)	<i>spe</i>  RS-5 (Glass 2007: 93)
<i>śpi</i>  BLS1 (Allon 2001: 59)	<i>smi</i>  BLS21 (Glass 2000: 174)
 BLS2 (Lenz 2010: 18)	<i>smṛ</i>  BL-20 (Glass 2000: 133)
<i>śpe</i>  RS-5 (Glass 2007: 93)	<i>hme</i>  Schøyen 1 (Glass 2000: 184)

The *kṣa* and *sta* characters are also included in the table to provide a more complete listing of the non semi-vowel clusters, and are highlighted in the shaded rows. It is noticeable that the unique ligatures used for these characters are simpler than the typical

¹⁴⁹ This version has been traced from images of the manuscript forwarded by Tim Lenz.

¹⁵⁰ This character has been traced from the image of the rubbing of the *ārā* inscription presented in Konow (1929: 162-165, Plate XXX).

combined characters. These clusters are commonly retained as reflexes of the original Sanskrit clusters (§§ 8.3.1.3, 8.3.3.1), and simpler written forms may have developed over time due to this common usage. At the same time, the other conjunct consonants have mostly been used in the record of conservative spellings, while the original Sanskrit clusters have been reduced to a more advanced form in most reflexes (§ 8.3). This suggests there was possibly less developmental pressure on these forms, which could be improvised as combined characters when they occasionally needed to be written.

6. Patterns of sound change

Knowledge of the sound changes that occurred in Gāndhārī is only accessible through the available orthographic record, which offers both a broad and in places more subtle representation of the assumed spoken sounds (Chapters 5 and 7). These changes can be understood and confirmed through comparison with common and well understood patterns of sound change recorded in the world's languages. The majority of changes which occurred in the consonants of the MIA languages, when compared with the OIA languages, involve common processes of assimilation and lenition. A brief overview of these types of sound change is provided in this section, intended as a reference point for analysing and categorising the developments which occurred in Gāndhārī, and their stages of development in comparison with other MIA languages.

6.1 Assimilation in consonant clusters

Assimilatory changes are very common in historical linguistics and occur when one or more properties of a sound are conditioned by and become more similar to another nearby sound. This normally occurs for reasons of articulatory ease, where the sharing of articulatory properties limits the necessary transitions of the speech articulators, although perceptual considerations may also play a role. Assimilation may be either regressive, in which the first consonant anticipates a feature or features of the following consonant, which spread backwards as in $C_1C_2 > C_2C_2$ ¹⁵¹, or progressive, in which a feature or features of the first consonant are carried forward into the following consonant, as in $C_1C_2 > C_1C_1$. Assimilation may also be total, in which a sound takes on all of the phonetic features of another sound, or partial, in which only some of the features are shared, so that the two sounds do not become identical. The spreading of place features is the most common form of assimilation on a cross linguistic basis, followed in frequency by the assimilation of voicing while the spread of manner features and total assimilation appear to be less common¹⁵².

¹⁵¹ These symbols indicate the progression between the first and second spoken sounds of the cluster, written in a left to right script. While this would be written in the reverse direction for a right to left script like Kharoṣṭhī, transliterated Gāndhārī text is typically presented using IAST symbols in a left to right script.

¹⁵² Campbell 2013: 24-25; Gordon 2016: 123-128

6.1.1 Perceptual and phonetic factors

The tendencies for assimilation to occur between different types of consonants in clusters has been evaluated in terms of the relative perceptual strength of their auditory cues, as well as some of their articulatory characteristics¹⁵³. Consonants rely on auditory cues¹⁵⁴ to establish their identity and provide a phonological contrast of manner and place with other consonants when perceived by listeners. These perceptual cues include both the internal auditory characteristics of the consonant, and cues that arise in the auditory transition between a consonant and any preceding or following vowel or consonant (Wright 2004: 36-41). It is suggested that segments which are encoded with greater redundancy, as well as sequences which result in a greater modulation of the acoustic signal, are more perceptually robust and likely to survive environmental degradation. These parameters are complementary and overlap to a reasonable extent, as for example in CV sequences, which benefit from the greatest perceptual overlap as well as optimal signal modulation. While CR sequences are only slightly less robust than CV, the stability of sequences such as CRVRC will depend upon the degree of separation between the auditory cues of the particular case vowel and approximant. In contrast, consonants with strong internal place and manner cues such as fricatives are more stable in the initial position, as in ST for example (Wright 2004: 42-50). This type of approach offers a useful tool in understanding some of the Gāndhārī developments, and a summary of the main auditory cum perceptual cues of consonants based on the manner of articulation is presented in Appendix B.

6.1.2 Place assimilation

The role of various perceptual and articulatory factors in place assimilation between consonants has been described in detail by Jun (1995, 1996, 2004). An important consideration is the relative position of the consonants in a $V_1C_1C_2V_2$ sequence, where the formant frequency transitions¹⁵⁵ and release burst that occur in the transition between C_2 and V_2 are typically stronger and of higher perceptual importance than the formant

¹⁵³ Jun 1995, 1996, 2004; Wright 2004; Gordon 2016: 129-132.

¹⁵⁴ A cue has been defined as “information in the acoustic signal that allows the listener to apprehend the existence of a phonological contrast” (Wright 2004: 36) while it is also noted that “the acoustic signal is produced by articulations that are continuous and overlapping to a greater or lesser degree” (Wright 2004: 36), and for this reason the resulting acoustic cues may also vary with context.

¹⁵⁵ Refer to Appendix B for an explanation of formants.

transitions between V₁ and C₁, which provide relatively weaker place cues¹⁵⁶. This suggests that both stops and nasals, which are dependent on these transitions for their identity, are stronger in the C₂ position and more exposed to place assimilation when placed as C₁. This is evidenced in the typological preference for regressive place assimilation on a cross linguistic basis, especially in clusters of stops and nasals, seen in the common example of place assimilation of nasals to following stops. Nasals in particular appear to be the most exposed due to the unreliable place cues of the nasal murmur formants, and the tendency for V₁ to become nasalised, which creates further confusion of place. In different combinations of nasals and stops, regressive place assimilation seems most likely to occur in nasal + stop, nasal + nasal and stop + stop combinations, while stop + nasal combinations may show mixed results due to the relative weakness of nasals compared to stops. In combinations of fricative + {stop, nasal} the place cues provided by internal frication are well preserved and appear to be less confusable, making fricatives more stable in the C₁ position, and similarly for approximants due to their internal formants and relatively slow formant transitions¹⁵⁷. Hence clusters of the type S{T, N} appear to possess an inherent stability since each element possesses relatively robust place cues, seen for example in Sanskrit clusters of the type *sk*, *sm*, *ʃp*, *śp* and *śm*.

The respective places of articulation of consonants also plays a role in their relative tendencies for regressive place assimilation in clusters, with coronals most likely to assimilate, followed by labials then velars which are least likely to assimilate¹⁵⁸. These tendencies are explained phonetically in terms of the relative size and speed of the articulators mobilized for different places of articulation, whereby the heavier tongue dorsum and lips in velars and labials have slower movements compared with the faster movements of the relatively small tongue blade articulator in coronals. This means that the formant transitions of velars and labials are likely to be more identifiable than those of coronals due to a longer transition time, while coronals are more prone to be overlapped by the slower sounds of the non-coronals. This suggest that regressive assimilation is more likely to be supported in a coronal + non-coronal cluster, while a non-coronal + coronal cluster will be more resistant, depending upon respective manner features. Hence neither

¹⁵⁶ Jun 2004: 61-62; Gordon 2016: 124, 128-129.

¹⁵⁷ Jun 2004: 62-63; Gordon 2016: 129-131.

¹⁵⁸ Jun 2004: 63-66; Gordon 2016: 131-132.

velars nor labials but only coronals are assimilated in many languages, as for example in the allophones *riḡh*[k] *call* for *right call* and *riḡh*[p] *pillow* for *right pillow*, heard in the pronunciation of many English speakers, while the velar and labial assimilation in *to*[k] *call* for *top call* and *so*[p] *pillow* for *sock pillow* occur less frequently (Gordon 2016: 131). It has also been suggested that velars are more resistant to assimilation than labials due to more compact formant transitions (Jun 2004: 64).

Exceptions of progressive place assimilation are seen in a number of languages, and the extent to which the above patterns apply consistently on a cross linguistic basis is unclear. For example, stems show a typological tendency to resist modification compared with affixes, leading to progressive assimilation of the initial consonant of suffixes in some languages (Jun 2004: 68)¹⁵⁹. Another example of progressive assimilation occurs between pairs of coronal consonants, where the second consonant mostly tends to assimilate to the first (Gordon 2016: 132). This occurs in Sanskrit coronal clusters involving the same manner of articulation but different places of coronal articulation, as in say retroflex plus dental or vice versa, seen in *aviḍḍhi* > *aviḍḍhi* and *ṣaṇṇām* > *ṣaṇṇām* for example. This is also seen in the forward spreading of retroflexion in many of the sandhi rules described in Appendix A1, including *T* > *Ṭ/ṣ* for example. It is suggested that the formant transitions into the closure from the preceding vowel are the main point of distinction between the dental [+anterior] and retroflex [-anterior] coronals (Gordon 2016: 132). This means that the first consonant has stronger perceptual clues, making the second consonant more prone to assimilation. This compares with the stronger perceptual cues to the second consonant in combinations that differ in the major class features of coronal, labial and velar. Alternate directions of assimilation between Sanskrit coronals are also seen, however, as in *tḥ* > *tḥ* in external sandhi for example, suggesting these developments could vary with different morphological processes, or relate to the spread of retroflexion.

6.1.3 Manner assimilation

While examples of place assimilation between consonants are widely reported in linguistic surveys, examples of manner assimilation appear to be far less common¹⁶⁰. At the same time manner features have been observed to spread less reliably¹⁶¹, and the

¹⁵⁹ Stems are usually stressed compared with affixes, which may be a reason for this phenomenon.

¹⁶⁰ E.g., Gordon 2016: 128.

¹⁶¹ E.g., Padgett 1994: 465.

patterns and motivations for manner assimilation appear to less well understood than for place assimilation.

One approach to explaining manner assimilation in medial consonant clusters is through hierarchies of strength, in which assimilation between consonants is analysed on the basis of the relative strength of their respective manners. This is typically defined in terms of the degree of oral stricture, broadly assigned in the sequence of stops, nasals, fricatives and semivowels, or $T > N > S > R$, where T is stronger than N and so forth. In this system the nasals are stronger than fricatives due to their closed oral stricture, while at the same time their nasally released acoustic signal is considered to be more sonorant (§ 4.4). Strength sequencing has been used to explain patterns of total assimilation observed in Pali for example, where, with one or two exceptions¹⁶², Sanskrit clusters typically develop as geminates in favour of the stronger of the two consonants, regardless of directionality, and involving concurrent assimilation of place in many cases¹⁶³. This is seen in Skt. *niṣka-* > P *nekkha-*, *karka-* > *kakka-*, *takra* > *takka-*, *kiṇva* > *kiṇṇa-*, *ramya* > *ramma*, *sarva* > *savva* and *arya* > *ayya* for example¹⁶⁴. The relative strengths of the semivowels have been further assigned in the order of *l*, *v*, *y* and *r*, although it is unclear whether this ordering has a clear phonetic basis or is assigned empirically according to the observed patterns of assimilation in Pali¹⁶⁵. In this strength-based approach weaker first consonants seem to anticipate the occlusion of stronger second consonants in regressive assimilation, while stronger first consonants spread their occlusion to weaker second consonants in progressive assimilation. This results in the formation of geminate consonants that are as strong as possible in the context of the original cluster, which is consistent with the tendency for less sonorant consonants to form geminates, as described in § 4.6. At the same time the phonetic distinction between consonants and adjacent vowels is maximised (Suzuki 2002b: 64-65), creating a greater difference in sonority as

¹⁶² Exceptions include Sanskrit stop-stop and nasal-nasal clusters which develop through regressive place assimilation to form geminates of the second element, and homorganic NT clusters between pairs of obstruents which remain stable. TN clusters develop as TT clusters however, according to the strength hierarchy, while *hR* clusters appear to undergo metathesis.

¹⁶³ E.g., Cho 1999: 167-172; Suzuki 2002a, 2002b; Oberlies 2019: 147-151.

¹⁶⁴ Suzuki 2002a 102-105.

¹⁶⁵ From a phonetic perspective these distinctions suggest that the semi-vowels can also be classified in terms of differences in manner as much as in place, which might be explained in terms of the subtle variations in the shape and placement of the tongue during articulation, as for example in the lateral shaping and slight central contact of the tongue during the articulation of *l*, the more open vowel like articulation of *v* and *y*, which is more open in the latter case, and the short tap of *r* in an otherwise open articulation (Davenport and Hannahs 2016: 31-35). On the other hand, this classification is contrary to the normal distinction between the liquids *l* and *r*, and the glides *y* and *v*, in terms of sonority and openness (§ 4.4).

well as a simpler and clearer speech pattern, but one in which there is less phonological contrast due to the loss of one of the consonants in the cluster. Another explanation of Pali manner assimilation, which derives from a phonetic interpretation of Sanskrit consonant doubling patterns (§ 4.6), is explored in § 9.3.

Converse patterns of assimilation, where the more sonorous elements dominate, have also been observed in other languages. For example, *tl* > *ll*, *ld* > *ll*, *ln* > *ll*, *nd* > *nn*, *rn* > *rr* and *mb* > *mm*, which include both regressive and progressive assimilation, are seen in some regional European and Italian dialects¹⁶⁶, while *ts* > *ss* is also reported¹⁶⁷. While these instances of gemination are consistent with the assumption of increased articulatory ease (Gordon 2016: 123-124), they also reduce the articulatory contrast between the geminate and surrounding vowels. These examples suggest that patterns of manner assimilation develop on a language specific basis, and approaches to manner assimilation based on sonority or strength hierarchies do not appear to offer a comprehensive cross linguistic explanation¹⁶⁸.

The typological preference for regressive place assimilation discussed in the previous section has led to the analysis of consonant clusters in terms of the relative strengths of the coda and onset in clusters having falling sonority. Known as Coda Target (Cho 1999: 101-105) or coda / onset asymmetry (McCarthy 2008: 271-273), in this approach the commonly observed regressive assimilation of coda consonants has been explained in terms of the phonetic weakness of consonants in syllable final position (Murray 1982: 172-173). Cho (1999: 101-105) has articulated this phenomenon to explain examples of place assimilation, and in these applications the assumed weakness of the coda is consistent with the relative weakness of the perceptual cues of initial consonant discussed in § 6.1.2, which can also be expressed in terms of the limited capacity of the coda to sustain place features (Gordon 2016: 106). Murray (1982: 172-173) has also analysed examples of regressive manner assimilation in Pali, such as Skt. *kar.ka* > P. *kak.ka*, in relation to coda weakness. While offering an explanation in situations where regressive manner assimilation occurs, it does not explain cases where the original cluster has rising sonority and forms a syllable onset, as in Skt. *ta.kra* > P. *tak.ka* for example. Such approaches can however be considered alongside other characteristics of the syllable

¹⁶⁶ Reported in Murray 1982: 168-169.

¹⁶⁷ Cho 1999: 170-172.

¹⁶⁸ E.g., Murray 1982; Cho 1999: 170-172; Wright 2004: 34-35.

boundary in a broader analysis of both regressive and progressive manner assimilation, and are discussed further in § 9.3.4.

The patterns of consonant cluster manner assimilation recorded for Pali are also evident in other MIA languages, although these developed to a lesser extent and with less consistency in Gāndhārī (§ 9.1.2). These divergences, which appear to represent intermediate stages of development, offer further insights into the progressive development of manner assimilation in different types of clusters, particularly with respect to assimilation of the semi-vowels.

6.1.4 Synchronic and diachronic considerations

It has been suggested by many scholars that the phonetic factors which drive assimilation and shape the phonology occur diachronically, through a progressive process of ‘innocent misapprehension’ (Gordon 2016: 133). In this model the two co-articulatory sounds become perceptually less distinct over time, and the one in a less noticeable position becomes vulnerable to re-analysis of the assimilating feature by listeners. In the alternate view it is suggested that speakers are aware at a synchronic level of the perceptual biases and constraints of the listeners, and reshape their phonologies accordingly.

6.2 Lenition, deletion and fortition

Lenition refers to the weakening of sounds, which is commonly assumed to involve a weakening of articulatory effort, which also relates to the environment in which a sound occurs. Changes in the strength of consonants typically relates to changes in manner as well as laryngeal features. In an articulatory based hierarchy, geminate stops are rated as strongest, followed by aspirated single stops, single stops, fricatives, approximants and then laryngeals, following a similar order to the strength hierarchy presented in § 6.1.3. Lenition of a consonant typically represents a loss of strength along this scale, which mostly involves a reduced articulatory duration or constriction, although voicing and aspiration features may also be involved. In all cases this should involve a reduced articulatory effort¹⁶⁹. Lenition is most commonly triggered by adjacent sounds which have a relatively open vocal tract, such as vowels or weaker and more open consonants. These

¹⁶⁹ Kirchner 2004; Campbell 2013: 37; Gordon 2016: 153-155.

may be placed on one or both sides of the target consonant. The weakening of single intervocalic stops is a common example of lenition. Deletion represents the full lenition of a sound, and essentially occurs in the same environment. Deletion may in turn lead to a deterioration of the syllable structure (Gordon 2016: 157).

Lenition encompasses some common classifications such as degemination, flapping, spirantisation and debuccalisation. Flapping involves the lenition of stops or trills into flaps, while spirantisation involves the lenition of stops into continuants, and debuccalisation typically involves the replacement of an oral consonant, commonly a voiceless fricative, with a glottal gesture such as aspiration, which effectively involves the loss or spreading of the place features to a glottal¹⁷⁰. While nasals can be produced with a variable width of the velic opening, it is observed that fricative nasals and nasal sounds that contrast in this way are generally not seen in the world's languages (Ladefoged and Maddieson 1996: 103).

Fortition refers to the strengthening of consonants, which typically involves a narrowing of the articulatory constriction, in contrast to the reduced constriction associated with lenition. This may be realised in changes such as the conversion of a fricative to a stop, or a glide to a fricative for example. Fortition appears to be far less common than lenition, and may be triggered by such factors as increasing the contrast between a consonant and a following vowel, whether through devoicing or increased constriction (Gordon 2016: 151-155).

¹⁷⁰ Fallon 2002: 123-127; Kirchner 2004: 313-314; Gordon 2016: 153.

7. Recovering the phonology of the Gāndhārī language

7.1 General methodology

The Gāndhārī language existed in a linguistic environment that is accessible today through a number of sources that provide significant information about its phonological system. These include:

- The growing body of written Gāndhārī material that is now available (Chapter 3).
- The written and oral records of the OIA and other MIA languages which have survived in to the present day, and which were part of the historic language environment surrounding Gāndhārī (§§ 2.2.2, 2.2.3).
- The well described grammar and phonology of Sanskrit (Chapter 4) and other MIA languages.
- The well understood general phonological changes that occurred between OIA and other MIA languages¹⁷¹.

Published descriptions of the Gāndhārī source materials typically include interpretations of the grammar and phonological system as represented in the materials, which has allowed a reasonable picture of the likely phonological system to be established (Chapter 5). All of the existing analyses are however published within wider studies of particular manuscripts or inscriptions, and there is some variability in how the phonemes and changes from OIA were recorded in the original sources, as well as in the classifications used in the published descriptions. While this body of work has provided an expanded data base and clearer understanding of the Gāndhārī phonological system, the material does not appear to have been collated and studied within a single volume, and there are benefits to be gained in collating this material for analysis in a more structured format.

The well documented Sanskrit grammar and phonology provides a clear reference point which is close to the languages from which Gāndhārī and the other MIA Prakrits have descended. While the MIA languages are less well documented than Sanskrit, the changes recorded in Gāndhārī typically follow similar or close patterns to those occurring in these languages, although with some exceptions such as the retention of all three

¹⁷¹ E.g., Pischel 1879; Misra 1967; Masica 1991; Oberlies 2019.

sibilants and limited extent of assimilation of the consonant clusters. When available, parallel inscriptions and texts have allowed word level comparisons to be made, which assists in identifying the phonological changes that occurred between OIA and Gāndhārī. The wider body of Aśokan rock inscriptions are recorded in Prakrit dialects, and carry a similar message to the Gāndhārī inscriptions at Shāhbāzgarhī and Mānsehra, allowing the phonologies of the different versions to be compared¹⁷². Parallel versions of the Gandhāran Buddhist manuscript texts have been identified in languages such as Pali, Prakrit, Buddhist Hybrid Sanskrit and standard Sanskrit. Where available, parallel or near to parallel sections of text are listed in all of the source publications¹⁷³, while there are no known parallels for the pūrvayoga and avadāna documents described by Lenz (2003: 79; 2010), nor the scholastic documents described by Schlosser (2016: 261). The Gāndhārī texts only partially follow the parallel versions in some instances, and it has been difficult for scholars to establish the historical relationship with particular known versions of these texts, meaning that the actual source document for the Gāndhārī text may differ from those in which parallel sections have been identified (Salomon 2014: 11-12). Many differences are also encountered between the parallel versions of early Buddhist manuscripts preserved in different languages and from different regions¹⁷⁴, adding further complexity to understanding the origins of the texts and source documents used in the preparation of particular Gāndhārī manuscripts.

Sanskrit is essentially a synthetic language, in which syntactic relations are largely expressed by inflectional morphemes, and word structures are built up with several layers of affixes added to the root in some cases. The boundaries of this morphological structure are veiled to some extent by the application of internal sandhi in the written texts (§ 4.2). When filtered through the changes which occurred in a language like Gāndhārī, and recorded in the Kharoṣṭhī script with its inherent limitations, the resultant morphological structure of the descendant language becomes quite opaque. The MIA languages were essentially vernacular dialects in which the OIA grammar and case systems had broken down considerably (§ 2.2.3), and the extent to which the grammatical structure was understood and articulated by scholars at the time is uncertain. The example of the Sanskrit causative verb *bhāvayati* which is formed as *bhava + (a)ya + ti* in Sanskrit,

¹⁷² E.g., Hultsch 1925: 185, Synoptical Texts.

¹⁷³ E.g., Salomon (2000: 204-205); Allon (2001: 307); Lenz (2003: 195); Glass (2007: 211); Salomon (2008: 413); Baums (2009: 687); Silverlock (2015: 592).

¹⁷⁴ These issues are discussed at length by Allon (2021: 47-108 [Chapter 5]).

provides a simple illustration of this point. This typically develops as *bhavedi* in Gāndhārī (GD s.v.)¹⁷⁵, which could simply be a record of the spoken sounds, or alternately could have been understood in a grammatical sense, as say *bhav(a) + e + di* or *bhava + i + ti* for example. Alternative reflexes such as *bhavae* and *bhavea* are also recorded in a small number of manuscripts (GD s.v.). While these variations suggest on the one hand that any formal understanding of grammatical structure was either limited or loosely applied, they may also be an indication that the language and its grammars were changing. Another factor which complicates analysis of the Gāndhārī texts is the lack of separating spaces between words in written documents, which are almost always in the Kharoṣṭhī script (§ 5.1). While this obscures the pattern of words in these texts, this is mostly overcome through comparison with parallel versions or familiarity with the morphological patterns of Sanskrit and MIA words.

7.2 Interpretation of orthographic intent

The phonological intention of the Kharoṣṭhī orthography is often unclear, and interpreting the phonemic system of the language from the written texts can be a complex task. As discussed in § 5, many diacritic markings which appear to indicate phonetic modification of the sounds have been added to the basic akṣaras in the Kharoṣṭhī script, and as noted by Strauch (2012: 148), this suggests that by its nature the script was phonetic rather than phonological:

“Throughout its history it showed a strong tendency to depict different allophones by different or modified letters. This resulted in a rather large diversity of signs which hardly correspond to the phonological structure of the language.” (Strauch 2012: 148).

In other words, it appears that the scribes may have been trying to capture the spoken allophones, using different or modified akṣaras. Fussman (1989 467, [§ 27.1]) has questioned whether the writers of Kharoṣṭhī would have been aware of the notion of the allophone, but this is not necessarily contradictory. It is possible the scribes were trying to adapt or modify the available script to what they heard, without fully understanding the nature of the phonetic variations, or being able to distill the most appropriate ‘underlying form’ with which to represent the actual phoneme. All phonemes are expressed as allophones, or surface speech sounds, which vary slightly in different situations, and the

¹⁷⁵ *bhavedi* is listed as the lemma.

phonetic magnitude of these changes may affect how they are perceived, seen for example in linguistic surveys which tend to underestimate the extent of assimilation that has occurred in spoken languages (Gordon 2016: 127). This highlights the progressive nature of conditioned sound changes, where intermediate stages occur in which the sound adjustments commence at the level of allophones, without any change in the word orthography, until the perception is strong enough to render the acceptance, or writing, of a new orthographic representation for the changed sounds. This does not necessarily involve changes to the phonemic inventory of the language, but rather the selection of alternate orthographic representations from within the inventory which are more appropriate to represent the changed sounds in particular environments. The extent to which such changes may develop in a language over shorter or longer timeframes is unclear, however (Gordon 2016: 133), although they typically involve an intermediate stage in which both the archaic and innovating forms coexist¹⁷⁶. In this context the Kharoṣṭhī script displays a diversity of both diachronic and synchronic orthographic usage (Strauch 2012: 148), whereby for example several alternate spellings might be found for one word. These are often used in an inconsistent manner, and can give the impression of a series of allographs used in free variation. These variants have been broadly classified in terms of:

- **Historic** or archaic forms, in which the conservative orthography is retained. These may reflect conservative administrative or scriptural usage, or borrowings from source documents, but may also reflect unchanged speech forms or just a desire to retain traditional spellings. These may also serve a structural purpose in the written language by distinguishing between homophones with different meanings. These may incorporate forms from earlier or different stages in the language development, as in English *rite* / *right* and *write* / *wright* for example.
- **Transitional** forms which are typically indicated by diacritic marks attached to the historic graphemes, but may take other forms in the case of consonant clusters for example. These appear to represent conditioned sounds, or sounds which are transitional between those of the historic and advanced forms. In the sense that these represent conditioned sounds they are most likely allophones of the historic forms.

¹⁷⁶ Weinreich, Labov and Herzog 1968: 188; Matsumoto 2019: 199.

- **Advanced** or innovating forms which appear to reflect full sound changes to the extent that these are now similar to those represented by other existing graphemes, which are used to record these reflexes. These represent a distributional change in the arrangement of the phonemes in particular environments. These forms involve a full shift in either the voicing, manner or place of a phoneme.

Variant forms are recorded for both single as well as clustered consonants in Gāndhārī, and reflect the sequential development of a series of different sound change processes. For example, in the case of single consonants, both the transitional and advanced forms are typically related to specific environments, such as V_V for example, while the original phoneme may be fully retained in word initial positions. Some simple examples of these patterns include:

- Intervocalic dental stops are commonly recorded with a diacritic marker on the original akṣara, as in $(k, g) > \underset{\cdot}{g} / V_V$ for example which has been interpreted to represent lenition to a fricativised pronunciation [ɣ]¹⁷⁷. In many cases both the modified and unmodified akṣaras have been used in the same lexeme, such as *loke* > *loga* / *loge* for example (§ 8.2.1.2.1), which also reflects the merger of *k* > *g* through voicing. This indicates that the scribes were aware of both the historical spellings as well as the changing nature of the spoken sounds.
- In addition to conservative and transitional forms, intervocalic reflexes of the velar stop *g* include use of the palatal approximant *y* on a reasonably regular basis, as in $g > y / V_V$, seen in Skt. *bhaga-* > G *bhaya-* for example. The recording of this reflex suggests on the one hand that the sound change has moved sufficiently in the direction of another recognised phoneme for its symbol to be used on a regular basis. On the other hand, since *bhaga-* is also recorded by some scribes, involving the transitional form *g*, the recording of *bhaya-* might also be interpreted as the use of an available and similar sounding grapheme to mark this allophone. These variants may also reflect synchronic dialectical variations.

Since these variable forms have been used inconsistently in many cases, it is possible they may have co-existed in the language over an extended period of time, while the use of advanced forms suggests that the extent of innovation in the language was quite well developed. On the other hand, the recording of transitional forms using diacritic

¹⁷⁷ E.g., Salomon 2008: 338; Baums 2009: 141-142; Silverlock 2015: 231-232.

marks suggest that the processes of sound change were still in progress, and some of the scribes attempted to capture these transitional allophones with unique diacritic markers. The use of advanced forms may in some cases have also related to the similarity of changed sounds with the those represented by available graphemes. The use of these variable forms suggests there was no established discipline in how to record the language, and the scribes were evidently aware of the existence of all of the phonetic variations, whether through variation in the natural speech, dialectical divergences, or other influences such as a loosely diglossic situation in the language. In this situation it is likely that the grapheme variants reflect the actual variation in the pronunciation of the language, and the ancient written record of the Gāndhārī language offers an insight into progressive stages of sound and language change that were alive at that time. Although it is not possible to measure the actual sounds of the language, these orthographic accounts provide an insight into the changing nature of the language at both a broad as well as more detailed level, and can be interpreted within the general articulatory pattern of the phonemes (§ 5). And while advanced forms may most closely represent the sound changes that occurred in the spoken language, the use of transitional forms suggests an inherent uncertainty, or even local variation in how the spoken language was interpreted, and how the written language should be presented. Some particular areas where advanced or vernacular spellings are more likely to have been used include the variable, infilled parts of formulaic inscriptions (Fussman 1989: 453 [§16.3]), the Buddhist pūrvayoga and avadānas texts, and scholastic texts such as BC 4, 6, 11 which may represent the notes of a student (§ 3.5).

7.3 Representation of Gāndhārī phonemes in the lemma forms of words

Single Gāndhārī consonants typically retain their original OIA form in word initial positions, and most of the variable reflexes for single consonants are recorded in word medial positions, where they are susceptible to intervocalic lenition. In these cases, it is appropriate to assess the most appropriate representation for each phoneme from the range of historic, transitional and advanced forms which are recorded in many cases. While the overall phonemic inventory of the language does not appear to have undergone significant change¹⁷⁸, potential changes to the distributional pattern of phonemes in words are suggested by the regular use of advanced forms, which are of particular interest in

¹⁷⁸ As represented in Tables 5.0, the main changes appear to be a loss of contrast between some of the nasals, and loss of the aspirated *jh* consonant from OIA (Table 4.0). These are addressed further in Section 8.2.

determining the phonemic forms to be used in the lemma forms of words¹⁷⁹. In most cases transitional forms appear to represent allophones of the historic forms, but assessments of the true extent of change represented by advanced forms presents as a more complex task. This is particularly so in cases where the graphemes of other consonants are used to represent advanced forms, suggesting a possible sound and distributional shift in the phonemes in particular environments. This can occur in several ways, such as the merger of phonemes through voicing, or the lenition of stops as approximants. This gives rise to a number of considerations:

- If other graphemes have been used to record advanced forms, it is possible these do not represent the true nature of the sound, but were rather a convenient way to represent the pronunciation of a phone which was close to that of the existing grapheme.
- While it might be assumed that the most phonologically advanced forms are closest to the vernacular language, if their use is limited in the written language, they may only be recent, dialectic or uncertain developments.
- The use of existing forms to represent changed sounds can lead to the development of homonyms, and a loss of phonemic contrast in minimal pairs. While these changes may have been accepted and understood in the language over time, the retention of traditional spellings may also reflect a desire to retain phonemic contrasts and avoid the development of homophones in the written language (§ 7.2).
- It might be expected that the orthography would have adapted slowly to the changing sound environment, and only adopt progressive forms once they were well established. In this context a full phonemic shift is most likely to have occurred when almost all of the reflexes are recorded in the advanced form.
- Non-etymological inverse spellings can also indicate merger between phonemes, as in the inter-mixed usage of *k* and *g* which appears to reflect merger as the voiced consonant *g* (§ 8.2).

The retention of historic spellings While the frequency of occurrence and stage of phonological advancement are important considerations, other factors need to be examined in determining the appropriate phonemic representation for a lemma. Consider as an

¹⁷⁹ The actual lemma may also change in a language record spanning some 600 years.

example the Sanskrit word *bhaga-vant*¹⁸⁰ (MW s.v. adj. “blessed one”), whose stem form has been represented in several ways in Gāndhārī, including *bhagava-* (the lemma form), *bhagava-*, *bhayava-*, and even *bhakava-* (GD s.v.). While the most advanced form of the stem is *bhayava-*, the frequency of use of *y* and *g* reflexes appear to be roughly similar in the overall record, while the transitional form *g* is used less often and mostly by one scribe (§ 8.2.1.2.1)¹⁸¹. Although a level of intervocalic weakening has clearly occurred, there is some uncertainty whether that has resulted in a full phonemic shift, and in this situation, it seems appropriate to retain the conservative form as the lemma. While a transitional representation such as *bhagava-* has the advantage of representing the change in progress, the diacritic marked *g* represents an allophone of the phoneme *g* in intervocalic environments, and its use would introduce an inappropriate and unnecessary phonetic form to the phonemic inventory.

Due to the variable phonemic record and inherent uncertainties in this process, it is proposed to establish a record of the main variant reflexes used for each phoneme, providing a background reference to the phonemic representations adopted in the lemmas. A shorthand method is proposed for this purpose, in which all of the variant reflexes are identified as /^h for historic, /^{tr} for modified or transitional, and /^{adv} for advanced, while further classifications of /^{alt} and /^{borr} can be used for alternate and borrowed forms respectively. This allows simple summaries of the variant forms to be prepared for each case. The following representation of the relationship between the lemma phoneme and its variant forms is suggested. The reflex of the Sanskrit suffix is enclosed in round brackets to indicate a notional form only.

G word: *bhaga-(va-)* (adj. “blessed”); Skt. *bhaga-vant*, P *bhaga-vanta*

$/g/ \equiv \{ /g^h, /g^{tr}, /y^{adv} \} / V_V$

In Gāndhārī *bhaga-(va-)* (adj. “blessed”) the phoneme */g/* in the stem has been represented by (\equiv) the historic phoneme */g^h/* and the allophones */g^{tr}/* and */y^{adv}/*.

It is beyond the scope of this thesis to analyse the full lexicon in this way, but by examining groups of words from a number of texts the most common reflexes for each Sanskrit phoneme can be recorded at a general level for both word initial and medial

¹⁸⁰ *Bhagavan*, the vocative form of this word, is often used as an eponym of the Buddha.

¹⁸¹ While the conservative spelling may have been retained more frequently in this word because the referent is the Buddha, and it is also parallel with the common Eastern spelling in Pali, there are many other words recorded in the manuscripts where the conservative -*g*- (or orthographic -*gh*-) spelling is retained.

positions. This allows the formulation of general rules to describe the development of each phoneme. The following format has been developed for this thesis¹⁸² and is proposed as a way of representing these generalised change rules. Forms that are less common can be bracketed, as notionally indicated for (/g^{tr}/) in the example.

Assumed sound change (Skt. > G): $g > g [\gamma] / V_V$
 $g > g [g] / \#_V$

Phonemic representation (G \equiv G): $/g/ \equiv \{ /g^h, /y^{adv} (/g^{tr}) \} / V_V$
 $/g/ \equiv /g^h / \#_V$

This format has been adopted for the representation of the variant Gāndhārī reflexes for both single consonants and consonant clusters, as described in sections § 8.2 and § 8.3. The possible use of advanced forms to represent some phonemic reflexes is also addressed further in § 8.2.1.2.6.2, and the selected phonemic representation can be readily adjusted in these formulae to suit new data or divergent opinion.

¹⁸² This format may have been used previously by others but has not come to the attention of the author of this work.

8. From Old Indo-Aryan to Gāndhārī: processes of phonological change

8.1 Introduction

The changes which occurred to the OIA consonants in Gāndhārī are examined in detail in this chapter. The changes which occurred to the single consonants are studied in § 8.2, while the changes to clusters of two or more consonants are analysed in § 8.3. The effects of these changes on the Sanskrit syllable structure are reviewed in § 8.4, on the retroflex consonants as a group in § 8.5, and on Sanskrit sandhi phenomena in § 8.6. While most of the analysis centres around changes recorded in the first and second century CE Gandhāran Buddhist manuscripts, which represent the largest collection of source materials, material from the Aśokan and non-Aśokan Kharoṣṭhī inscriptions is also included as this provides an important diachronic record beginning in the third century BCE. A number of unusual developments which are recorded in later stage source materials from around the third century CE are also examined in § 8.7. The following comments apply to the material tabulated in § 8.2 and § 8.3 unless noted otherwise:

- The Gāndhārī forms are reproduced from tabulations and examples included in the published source materials. In some cases, these are listed in stem form, while in others the fully inflected words from the Kharoṣṭhī text are listed. The Gāndhārī forms are typically copied as written in the published material, although inflected forms are reduced to stems in cases where the formation of the stem is clear.
- Sanskrit equivalents are mostly listed alongside the Gāndhārī examples in the published documents, but where these are absent in the text, they have been obtained from word index listings in the relevant document.
- A tabulation of Sanskrit consonant clusters, developed around manner and place of articulation, is presented in Table A3.1 in Appendix A3. This provides the basis for assessing the extent of the representation of these clusters in the Gāndhārī sources.
- The source documents are identified by the abbreviated ‘tags’ developed in Table 3.6. These have also been listed in Table 8.1 for ease of reference, including page numbers for the relevant sections. The standard abbreviations for the published Buddhist texts are also listed where relevant.
- The Eras are represented as follows: 3 BCE = third century BCE, 2 CE = second century CE, and so forth.

While the tabulations of examples are not comprehensive, the aim has been to capture at least one example from each source in which a reflex has been observed. A higher number of listings generally indicates that a reflex is recorded in more of the source documents. While this may also reflect more commonly used words in some cases, it does provide an indication of the extent to which a reflex may have been adopted or spread. Where variant reflexes are recorded for the same consonant or cluster, these are all listed.

Table 8.1: Source document reference abbreviations (Abbr.) and texts

Abbr. & Text	Reference	Abbr. & Text	Reference
Aś-H	Hultsch (1925: lxxxvii-xc)	BCS5-1	Schlosser (2016: 72-90)
Aś-M	Mehendale (1948: 10-27, 210-237)	BC4, 11	
InsK-K	Konow (1929: xcvi-cxi)	BCS5-2	Schlosser (2022: 78-97) ¹⁸³
InsK-M	Mehendale (1948: 305-314)	BC 4, 6, 11	(sd)
InsK-F	Fussman (1989)		
BLS1-1 AG-G ^L	Salomon (2008: 107-127)	BLS9-1 Khvs-G	Salomon (2000: 81-92)
BLS1-2 EĀ-G	Allon (2001: 79-98)	RS-5 ¹⁸⁴ SĀ-G	Glass (2007: 113-124)
BLS1-3 Dhp-G ^L	Lenz (2003: 40-45)	RS-12 Gos-G	Silverlock (2015: 229-273)
BLS2-1 PY-G	Lenz (2003: 120-135) (py)	RS-14 AG-G ^S	Salomon (2008: 337-342)
BLS2-2 Av-G	Lenz (2010: 25-36) (av)	RS-19	Lee (2009: 54-58)
BLS4-1 Nid-G	Baums (2009: 137-188)	RS-20	Marino (2017: 116-133)

A level of linguistic analysis and categorisation of the changes has been included in many of the publications describing the source materials. The works of Brough (1962), Fussman (1989), Salomon (2000, 2008), Baums (2009) and Silverlock (2015) are the most

¹⁸³ Additional material describing the BC fragment 6, which was not incorporated in Schlosser (2016), is included in this document. Due to the shorter length of fragment 6 compared with fragments 4 and 11, this only describes a limited number of distinct additional reflexes which were not described in the earlier work.

¹⁸⁴ Although the RS manuscripts were all written by the same scribe (§ 3.5), the distinction between manuscripts is maintained in this analysis in order to observe and trace any variations that may occur between the reflexes recorded in individual manuscripts.

comprehensive in this regard, and have provided important points of reference for the work in this section. In cases where the source genres and texts are metrical, the Gāndhārī translators generally appeared to have been more interested in accurate phonological transposition rather than metrical precision, and any irregular or unusual metrical patterns that arose appear to have been mostly accepted by the translators¹⁸⁵.

8.1.1 The ordering of consonants in tabulations and charts

The ordering of sections in § 8.2 and § 8.3 varies slightly from established conventions in the ordering of consonants in both Sanskrit and linguistics scholarship. The importance of strength hierarchies based on manner is well established in understanding the assimilation of consonant clusters in the MIA languages (§ 6.1.3), and for this reason material is primarily dealt with in the order of the manner hierarchy of stops, nasals, fricatives and approximants. While linguistic textbooks mostly follow this ordering, which parallels the gradual lessening of oral stricture implicit within the sequence¹⁸⁶, this varies from normal Sanskrit treatments which typically deal with the semi-vowel approximants ahead of the fricatives¹⁸⁷. Furthermore, while the Sanskrit semi-vowels are normally sequenced as *y, r, l, v* according to their places of articulation, they are sequenced as *l, v, y, r* within the manner strength hierarchies (§ 6.1.3), and this approach has been adopted to be consistent with the primary ordering according to these hierarchies. Where appropriate, sub-sections based on the place of articulation are included within each manner grouping, and these follow the Sanskrit *varṇamālā* sequence of velars, palatals, retroflexes, dentals and labials. While this is the converse of the common linguistic sequence¹⁸⁸, it provides an easier point of reference for readers familiar with this system.

On the other hand, when charts of the consonants are presented, these are organised in accordance with the established linguistic conventions of the IPA chart in order to present the OIA and MIA phonologies within a framework familiar to linguists¹⁸⁹. While many of the consonant chart tabulations in the Sanskrit literature are organised in a similar way, these differ in the ordering of the categories as mentioned above.

¹⁸⁵ E.g., Brough 1962: 79; Salomon 2000: 49-51; Lenz 2003: 24-29; Salomon 2008: 171-172.

¹⁸⁶ E.g., Ladefoged and Maddieson 1996: v-vi; Davenport and Hannahs 2010: 36-37.

¹⁸⁷ E.g., Whitney 1924: 2-3; Macdonell 1927: 4; Allen 1953: 20; Burrow 1955: 67; Coulson 1976: 8-13.

¹⁸⁸ E.g., Ladefoged and Maddieson 1996: 15; Davenport and Hannahs 2010: 13.

¹⁸⁹ See for example Tables 4.1, 5.0, 8.2.1.2.6a, 8.2.1.2.6b and 8.2.6.

8.1.2 The recording of geminates which are reflexes of consonant clusters

Many of the Sanskrit consonant clusters are assimilated as geminates in the Gāndhārī reflexes, and these geminates have been recorded in two ways in the published literature. Although the authors of earlier publications such as Hultsch (1925) and Konow (1929) have commented on the assumed gemination, they have not indicated it as such in the transliterations. On the other hand, the authors of the GBT series and other contemporary publications have indicated the assumed gemination by the writing of a second bracketed consonant in the tabulations of reflexes, as in $bd > d(d)$ for example. However, the bracketing is alternately applied to the first consonant in cases when there is an aspirated second consonant is the original cluster, as in $bdh > (d)dh$ for example. Since both regressive and progressive assimilation patterns occur in the assimilation of Sanskrit medial consonant clusters in Gāndhārī (§ 8.3.6.2), a convention of bracketing the assimilated consonant in the geminate reflexes has been adopted in all of the relevant tabulations. This reflects the observed assimilation patterns and directions of linguistic influence between the consonants in the original cluster. Hence the above examples would be represented as $bd > (d)d$ and $bdh > (d)dh$.

8.2 Changes to uncombined OIA consonants in Gāndhārī

8.2.1 Stops

8.2.1.1 Word initial

OIA word initial stops generally remain stable in Gāndhārī¹⁹⁰. A number of atypical developments have been noted however, which include the following:

- The writing of *gh-* for *g-* by the BLscribe 1, seen in Skt. *grāmam* > G *ghramu*, which appears to be scribal habit in both medial and initial positions¹⁹¹.
- A similar recording of aspiration in BLS2-2 appears to be the result of aspirate shifting, seen in some forms of OIA $\sqrt{\text{grah}}$, such as *grhnatī* > *ghin[ati]* where *ṛ* develops as *i*, while the aspirate position is preserved in other cases where the *ṛ* develops as *ra*, as in *grhītāḥ* > *grahida*, (Lenz 2010: 27-28).
- The writing of *dh-* for *d-*, seen in *dakṣinā* > *dhakṣiṇa* for example, is reasonably common and appears to represent a dialectical weakening of the distinction between aspirates and non-aspirates among the scribes of the Gāndhārī texts¹⁹².
- The development of *paruśa-* > *pharuśa-* (Salomon 2008: 107; Baums 2009: 137), which may be due to borrowings from Pali.
- The de-occlusion of *kh-* > *h-* which is commonly seen in *khalu* > *ho*¹⁹³.
- The de-occlusion of *bh-* > *h-* in *bhūtaḥ* > *hoda-* reflects a change seen in many word medial examples (Baums 2009: 138). Refer also to § 8.2.1.2.5.
- The alternation of *v* and *b* in some *b* reflexes, seen in *ucyate* > *bucadi*, *vucadi* for example. This is not uncommon in Gāndhārī (Allon 2001: 78).
- The treatment of enclitics as intervocalic in many cases, as in *tatra ca* > *tatra ya* for example, but not for example in *tena caiva* > *teṇa ceva*¹⁹⁴.
- The development of *t* > *c* in stem *tiṣṭha-* > *ciṭha-* appears to be a borrowed form from another MIA dialect, where this change occurs commonly (Baums 2009: 138).

¹⁹⁰ E.g., Mehendale 1948: 10-21, 297-305; Allon 2001: 78-79; Glass 2007: 113-114; Salomon 2008: 106-107; Baums 2009: 136-137; Lee: 2009: 53; Silverlock 2015: 222-223; Schlosser 2016: 71, Marino 2017: 116-117.

¹⁹¹ Allon 2001: 78; Salomon 2008: 106.

¹⁹² Allon 2001: 78; Lenz 2003: 42; Salomon 2008: 107.

¹⁹³ Glass 2007: 115; Salomon 2008: 109; Silverlock 2015: 232; Schlosser 2016: 71.

¹⁹⁴ Allon 2001: 78, 81; Marino 2017: 119.

- The alternative spelling of *dhyāna* > *jana-* or *caṇa* in RS-12 (Silverlock 2015: 222-223) appears to reflect confusion on the part of the RS scribe concerning the pronunciation of these two palatals, more commonly reflected in the medial cases (§ 8.2.1.2). It has been suggested that this indicates a merger between the pronunciation of these sounds¹⁹⁵. Another example is seen in *cittasya* > *jitaṣa* and *cittam* > *cito* in RS-5 by the same scribe (Glass 2007: 115), where the same word is spelt both ways.

The treatment of word initial stops in compounds is variable in Gāndhārī. Initial members of OIA compounds typically reflect the stem forms, which are vowel final in most cases, while following word initial stops remain unchanged in sandhi. While this format is mostly retained in Gāndhārī, they are treated as intervocalic medials in some cases and adjusted accordingly (§ 8.2.1.2). This is seen for example in Skt. *strī-kuṇapam* > *G istri-kuṇavu* and *varśa-kotīnām* > *barśa-kuḍiṇa* compared with *ātma-kāmena* > *[atva]-ghameṇa* and *citra-kathaḥ* > *citra-ghasu* (Salomon 2008: 106), and *khaḍga-viṣāṇa-kalpo* > *kharga-visaṇa-gapo* (Salomon 2000: 82, 229). The use of unmodified spellings suggests an awareness and retention of the traditional orthography (Silverlock 2015: 223-224), possibly in order to reflect the structure of these compounds, while the use of modified spelling suggests an awareness of the lenition of these sounds in the spoken language.

Summary: OIA word initial stops generally remain stable in Gāndhārī, and the development from Skt. > G is typically represented as /T-/ > /T-/. Some exceptions and occasional alternate representations do occur, as noted in the above bullet points.

8.2.1.2 Word medial

The Gāndhārī reflexes of OIA intervocalic single stops recorded in the Aśokan inscriptions and Gandhāran Buddhist manuscripts are presented in Tables 8.2.1.2a and 8.2.1.2b. The developments vary for each place of articulation, while the aspirated and unaspirated stops also develop differently. Variable reflexes are recorded in many cases, with divergent reflexes also recorded by the BL and RS scribes. Reflexes are ordered according to relative frequency based on an overview of the sources, while those which appear to be irregular or occur on an infrequent basis are enclosed in round brackets. The developments for each group of stops are discussed in the following sub-sections, while the overall patterns of development are assessed in § 8.2.1.2.6, including discussions regarding appropriate forms of phonemic representation.

¹⁹⁵ Glass 2007: 115; Silverlock 2015: 223.

Table 8.2.1.2a: Gāndhārī reflexes of OIA intervocalic stops – Aśokan and BL sources

Era	3 BCE	1 CE					
Source:	Aś-H Aś-M	BLS1-1	BLS1-2	BLS1-3	BLS2-1	BLS2-2	BLS4-1
pages:	lxxxv 210-214	107-116	79-84	40-42	128-130	25-28	137-147
OIA -k-	<i>k, (y)</i>	<i>Ø, gh, (k), (h)</i>	<i>Ø, gh, k</i>		<i>g</i>	<i>g</i>	<i>g, y, Ø</i>
-kh-		<i>h,</i>				<i>h (p 28)</i>	<i>h, Ø</i>
-g-	<i>g, k, y</i>	<i>gh, Ø</i>	<i>gh, y</i>	<i>gh</i>	<i>g</i>		<i>g, Ø</i>
-gh-	<i>h</i>	<i>h</i>		<i>h</i>			<i>h, kh, Ø</i>
-c-	<i>c</i>	<i>y, Ø</i>	<i>y, Ø</i>	<i>y</i>			<i>y, (g), (Ø)</i>
-ch-							
-j-	<i>j, y, c</i>	<i>y, Ø</i>	<i>j uncertain</i>	<i>Ø</i>	<i>y, (j)</i>	<i>y, j, c</i>	<i>y, (Ø)</i>
-jh-							
-ṭ-	<i>ṭ</i>	<i>ḍ, (ṭ)</i>				<i>ḍ</i>	<i>ḍ, ṭ</i>
-ṭh-							
-ḍ-			<i>ḍ</i>				<i>ḍ</i>
-dh-			<i>ḍh</i>				<i>ḍh</i>
-t-	<i>t, d</i>	<i>d, (t), (ḍ)</i>	<i>d</i>	<i>t, d, Ø</i>	<i>d</i>	<i>d</i>	<i>d, ḍ</i>
-th-		<i>s</i>	<i>s</i>	<i>s, h</i>	<i>ś, th</i>	<i>ś</i>	<i>s, ḍh</i>
-d-	<i>d</i>	<i>d</i>	<i>d</i>	<i>Ø</i>			<i>d</i>
-dh-	<i>dh</i>	<i>s</i>	<i>s</i>		<i>ś, s</i>	<i>s, ś</i>	<i>s</i>
-p-	<i>p, v</i>	<i>v, Ø, p, (b)</i>	<i>v, Ø, p</i>		<i>v, p</i>	<i>v</i>	<i>v, Ø</i>
-ph-							
-b-		<i>(Ø), (v)</i>				<i>v, Ø</i>	<i>v</i>
-bh-	<i>bh, h</i>	<i>h, bh, (Ø)</i>	<i>bh</i>				<i>h, v</i>
Reflexes in the table are listed in their apparent order of frequency. In this table round brackets () indicate reflexes which occur in limited numbers or are atypical.							

Table 8.2.1.2b: Gāndhārī reflexes of OIA intervocalic stops – BL, BC and RS sources

Era	1 CE	2 CE					
Source:	BLS9-1	BCS5-1 (BCS5-2)	RS-5	RS-12	RS-14	RS-19	RS-20
pages:	81-86	72-76	113-117	229-240	337-340	54-55	116-124
OIA -k-	<i>g, Ø, (k)</i>	<i>ḡ, Ø</i>	<i>Ø, k, k̲, g, ḡ, h</i>	<i>ḡ, Ø, (k), (y), (h)</i>	<i>k, k̲, Ø</i>	<i>Ø</i>	<i>Ø, k, k̲, g, ḡ</i>
-kh-	<i>(h)</i>	<i>k, h, Ø</i>	<i>h, g</i>	<i>h</i>	<i>g, kh,</i>	<i>h</i>	<i>h</i>
-g-	<i>g, (Ø)</i>	<i>ḡ, y, Ø</i>	<i>g, ḡ, y, k̲</i>	<i>ḡ, g, y,</i>	<i>Ø</i>	<i>k̲, y</i>	<i>Ø, y, k, k̲, g</i>
-gh-	<i>(h)</i>	<i>h ((k̲h))</i>		<i>g, ḡ, (Ø)</i>	<i>g</i>		<i>g</i>
-c-	<i>y</i>	<i>y</i>	<i>y</i>	<i>y, Ø, (c)</i>	<i>Ø</i>	<i>y, j, Ø</i>	
-ch-							
-j-	<i>y</i>	<i>y, Ø</i>	<i>y, j, c</i>	<i>y</i>	<i>y</i>	<i>y</i>	<i>y, c</i>
-jh-							
-ṭ-		<i>ḍ</i>	<i>ḍ</i>	<i>ḍ</i>	<i>ḍ</i>		<i>ḍ</i>
-th-	<i>(ḍh)</i>						
-ḍ-	<i>(ḍ)</i>						
-dh-		<i>ḍ (?)</i>					
-t-	<i>t, d, (ḍ), (p 83)</i>	<i>d, (ḍ), (s), (Ø)</i>	<i>ḍ, d, t/d, (s, ḡ, Ø)</i>	<i>ḍ, (ḍ)</i>	<i>ḍ</i>	<i>ḍ, (Ø)</i>	<i>ḍ, d, t/d,</i>
-th-	<i>ḡ, (s)</i>	<i>ḡ, (ḍh)</i>	<i>ḡ, s</i>	<i>s, ḡ, (ḍ)</i>	<i>s, ḡ</i>	<i>s</i>	<i>s, ḡ</i>
-d-	<i>t, (d)</i>	<i>d</i>	<i>ḍ, t/d</i>	<i>ḍ, (d)</i>	<i>ḍ</i>	<i>ḍ, (Ø)</i>	<i>ḍ, d, (d)</i>
-dh-	<i>ḡ</i>	<i>ḡ</i>	<i>ḡ, s, dh</i>	<i>s, ḡ, (dh)</i>	<i>s</i>	<i>ḡ, s</i>	<i>s, ḡ</i>
-p-	<i>v, (Ø)</i>	<i>v, Ø</i>	<i>v, Ø</i>	<i>v, Ø, (p)</i>	<i>v</i>	<i>v, Ø</i>	<i>p, v, Ø</i>
-ph-							
-b-	<i>b, (v)</i>	<i>((v))</i>		<i>(v)</i>			
-bh-	<i>h, bh</i>	<i>bh, vh, h</i>	<i>v, bh, Ø</i>	<i>v, (bh), (y)</i>	<i>v, h</i>	<i>v, h, Ø</i>	<i>bh, v, h</i>
Reflexes are listed in their apparent order of frequency. In this table round brackets () indicate reflexes which occur in limited numbers or are atypical.							

8.2.1.2.1 Velar stops

The Gāndhārī reflexes of OIA single intervocalic stops recorded in the source materials are summarised below, expressed in the format of Skt. consonant > G reflexes.

Unaspirated voiceless velar stop:

Aś Ins: $k > \{k, (y)\} / V_V$

BL Scribes: $k > \{g \text{ (gh)}, \emptyset, (k), (h), (y)\} / V_V$

RS & BCS5 Scribes: $k > \{g, \emptyset, k, \underline{k}, g, (h), (y)\} / V_V$

Unaspirated voiced velar stop:

Aś Ins: $g > \{g, k, y\} / V_V$

BL Scribes: $g > \{g \text{ (gh)}, \emptyset\} / V_V$

RS & BCS5 Scribes: $g > \{g, \underline{g}, \emptyset, y, \underline{k}, (k)\} / V_V$

Some of the examples given in the source texts are:

Source	Skt.	Sanskrit > Gāndhārī
Aś Ins.	<i>k</i>	$k > y$: <i>ardhatrika-</i> > <i>adhātiya-</i> (Suffixal only. Aś-M p. 11)
	<i>g</i>	$g > k$: <i>Maga-</i> > <i>Maka-</i> , <i>-upaga-</i> > <i>-upaka-</i> (Aś-M p. 11) $g > y$: <i>-upaga-</i> > <i>-upaya-</i>
BLS1-1	<i>k</i>	$k > k$: <i>ekam</i> > <i>eka</i> $k > gh (g)$: <i>ātmakāmena</i> > <i>[atva]ghameṇa</i> , <i>janakāyaḥ</i> > <i>jaṇaghayū</i> $k > \emptyset$: <i>anadhika-</i> > <i>aṇasia-</i> , <i>anukampayā</i> > <i>aṇoapae</i> , <i>pāpikām</i> > <i>pavia</i>
	<i>g</i>	$g > gh (g)$: <i>ārāgitaḥ</i> > <i>araghidu</i> , <i>kālagataḥ</i> > <i>kalaghadu</i> , <i>sugatiśu</i> > <i>sughadiśu</i> $g > \emptyset$: BHS <i>antraguṇam</i> > <i>(*a)traūṇa</i> , <i>bhaginī</i> > <i>bhaiṇi</i>
BLS1-2	<i>k</i>	$k > k$: <i>ekam</i> > <i>eka</i> , <i>akuśalāḥ</i> > <i>akuśala</i> $k > gh (g)$: <i>prakāśet</i> > <i>praghase</i> , <i>loka-</i> > <i>logha-</i> $k > \emptyset$: <i>adhikaraṇam</i> > <i>asiaraṇam</i> , <i>viveka-</i> > <i>vivea-</i>
	<i>g</i>	$g > gh$: <i>virāga-</i> > <i>viragha-</i> , <i>sugataḥ</i> > <i>sugha[du]</i> $g > y$: <i>bhagavant</i> > <i>bhayavada</i>
BLS1-3	<i>g</i>	$g > gh (g)$: <i>P āgamanāya</i> > <i>[a]ghamaṇa</i> , <i>uragaḥ</i> > <i>uragha</i>
BLS2-1	<i>k</i>	$k > g$: <i>upakaraṇam</i> > <i>upagarano</i> , <i>dārakāḥ</i> > <i>daraga</i>
	<i>g</i>	$g > g$: <i>āgacchati</i> > <i>agachadi</i> , <i>āgatam</i> > <i>agada</i> , <i>nagare</i> > <i>nagare</i>

BLS2-2	k	k > g: ājivik-a > āyiviga-, jihonikaḥ > jihonige, sthavikāḥ > thaviga
BLS4-1	k	k > g: akāmam > agamo, *ākāśa- > agaśa-, prākṛtikāḥ > pragidia, udakam > udaga, aśokaḥ > aśogo, lokaḥ > loga, kṛtakāḥ > kridaga, *kalyāṇaka- > [ka]ḷaṇaga- k > Ø: audārikānām > oda[r]iaṇa, pārāyaṇikaḥ > pa[r]ayaṇio,
	g	g > g: rāgaḥ > raga, g > Ø: prayoga- > prayoa-
BLS9-1	k	k > g: nipakam > ṇivago k > Ø: sāmāyikam > samaia
	g	g > g: mṛgaḥ > mṛigo, nāgaḥ > ṇago, rāgam > raga
BCS5-1	k	k > g: akārā- > aḡara-, viveka- > viveḡa- k > Ø: kapālakena > kavalaēṇa, caitasika- > cedaṣia-
	g	g > g: anāgata- > aṇagada-, vairāga- > veraḡa- g > y: kāmabhogin > kamabhoyi g > Ø: paligodhena > [a]palioṣeṇa, suvairāga- > suverao-
RS-5	k	k > k: eka- > eka-, ekagrata > ekagraḍa, pratikūla- > paḍikula-, parikarṣati > parikaṣaḍi, k > g: loke > loga k > g: loke > loge k > h: yuśmākam > tuspahu k > Ø: abhyavakāśa > abhoaśa
	g	g > g: a-nagaram > aṇagara; g > g: saḡagata- > saḡagada- g > y: bhagavān > bhayava g > k: āḡārāt > aḡaraspa, bhāvanānuyogam > bhavaṇaṇuyoka
RS-12	k	k > k: akuśalaiḥ > akuśal[e] k > g: lokasya > logasa, lokaḥ > loge, ākāśaḥ > [aḡa]śa k > Ø: kāya-karman- > kayiame-, nandikaḥ > ṇadio, pāpakaiḥ > pavaehi
	g	g > {g, ḡ, y}: bhagavant > bhagava / bhagava / bhayava
RS-14	k	k > k: śrāvakaḥ > śavaka, pratyekabuddha > pracekabhudasa k > k: vipākāḥ > vivaḡe, śrāvaka- > śavaḡa- k > Ø: anarthikaḥ > aṇathio, śaravaka > śavao
	g	g > Ø: ārāghitaḥ > araḡde
RS-19	k	k > Ø: -antike > -ṣatia, -nikāyam > -ṇiao, ādhyātmikānām > aḡatviaṇa
	g	g > k: aḡāra- > aḡara-, -rāga- > raḡaṣ(a)- g > y: bhagavantam > bhayavata
RS-20	k	k > k: ekam > eka

		<i>k > <u>k</u>: śoka > śo<u>k</u>a</i> <i>k > g: loke > logo</i> <i>k > <u>g</u>: loke > loga</i> <i>k > Ø: upāsakam > uasao</i>
	<i>g</i>	<i>g > g: yogah > yoge</i> <i>g > k: nagarasya > ṇakarasa</i> <i>g > <u>k</u>: nagaram > ṇa<u>k</u>ar[e], -gāminī > -<u>k</u>amiṇa (compound final)</i> <i>g > y: bhagavān > bhayava</i> <i>g > Ø: ayo-<u>g</u>uḍah > aya-üḍa, *mṛ<u>g</u>a > mia</i>

A variety of developments for *k* and *g* are recorded in the manuscript sources, which appear to be used arbitrarily both within and between documents, as well as among the different scribes, including in the Aśokan inscriptions. While this makes it difficult to discern any clear patterns of use, some reflexes tend to be used more frequently in particular situations than others.





Original intervocalic *k* is sometimes retained when it follows the negating prefix *a-*, seen in Skt. *akuśalāḥ* > G *akuśala* (BLS1-2) for example, which is similar to reflexes seen in the Dhp-G^K (Allon 2001: 80). It is also normally retained in *ekam* > *eka* where it appears to follow the common MIA reflex¹⁹⁶. The most common reflex is intervocalic voicing as *g*, which is a form of assimilatory lenition since less articulatory effort is required to maintain continuity of voicing with adjacent vowels (Gordon 2016: 152). Since original *g* is retained in many cases, as in *anagaram* > *aṇagara* (RS-5) for example, the record suggests that intervocalic *k* merged with *g* at some stage in the language development (Baums 2009: 140), while the occasional inverse spelling of *k* for *g* also supports this assessment. The character *gh* is often written for both *k* and *g* in what appears to be a graphic habit of the BL scribe 1, used for example to represent a *g* reflex at compound boundaries, or in derivations from the root $\sqrt{kāś}$ ¹⁹⁷. While the aspiration of pre-vocalic voiceless stops is a common phonetic occurrence in speech, this is not the case for voiced stops¹⁹⁸, supporting an interpretation that the writing of *gh* is a graphic habit rather than a record of allophonic aspiration. On the other hand, it is possible this was used to mark a fricative pronunciation, indicated by the fricative release of *gh*, similar to the use of diacritic marked reflexes discussed below.

¹⁹⁶ Salomon 2000: 82; Marino 2017: 117-118.

¹⁹⁷ Allon 2001: 80; Salomon 2008: 106.

¹⁹⁸ Ladefoged and Maddieson 1996: 66-70; Davenport and Hannahs 2010: 22-23, 69-70.

The other common reflex of original *k*, as well as for *g*, is elision (\emptyset). This is commonly seen in palatal environments involving the vowel *i*, as in Skt. *pāpikām* > G *pavia* (BLS1-1) and *samayikam* > *samaia* (BLS9-1) for example, although it also occurs in non-palatal environments, such as in *abhyavakāśa* > *abhoāśa* (RS-5) and *suvaīāga-* > *suverao-* (BCS5-1). It has also been observed that suffixal *-ka* appears to be especially subject to elision in Gāndhārī manuscripts¹⁹⁹, as in *caitasika-* > *cedaśia-* (BCS5-1) and *śaravaka* > *śavao* (RS-14) for example, although lenition also occurs in many suffixal cases, such as *nipakam* > *ṇivago* (BLS9-1) and *śrāvaka-* > *śavaḥa-* (RS-14). Where *k* and *g* are elided vowel hiatus is typically retained, and adjective endings in *-aka* are regularly recorded as *-ea*²⁰⁰. The lenition of *k* > *y* is recorded less often than the elision of *k* > \emptyset , although the development of *g* > *y* appears to be more regular. Along with the retention of vowel hiatus the recording of *y* suggests the consonant was still sounded, although in a significantly weakened form²⁰¹. The writing of *y* as the most advanced retained form suggests lenition as an approximant, while elision suggests weakened pronunciation as a barely audible *ya-śruti* serving as a syllable separator (e.g., Misra 1967: 166). This was possibly realised phonetically in the vowel hiatus, and it appears that the scribes did not consider it necessary to record this, consistent with the scribal laxity which is often observed in Gāndhārī documents (Salomon 2008: 338). This could also be interpreted to suggest that total loss was occurring in some cases.

In addition to *y* and \emptyset , reflexes of *k* and *g* are recorded by the RS scribe for both *k* and *g*. These are marked in the script with diacritic attachments to the original consonant graphemes seen for example in *ka* () and *ka* (), and *ga* () and *ga* () written by the RS scribe (Glass 2007: 91, 93)²⁰². It has been suggested that *ga* may have been pronounced as the voiced velar fricative [ɣ], representing an intermediate stage of lenition in the sequence *k* > *g* > *g* > *y* / \emptyset .²⁰³ These reflexes sometimes alternate with others in the same manuscript, seen for example in the variable spellings of Skt. *bhagavant* as G *bhagava*, *bhagava* and *bhayava* by the RS scribe in the Gos-G (Silverlock 2015: 231-232). While the alternation of historic *k* or *g* with advanced spellings such as *y* or \emptyset may be explained in terms of a desire to represent traditional forms in a conservative scholarly

¹⁹⁹ Salomon 2000: 82; Allon 2001: 80; Salomon 2008: 108-109.

²⁰⁰ E.g., Salomon 2008: 108; Schlosser 2016: 73.

²⁰¹ E.g., Mehendale 1948: 11; Salomon 2000: 82; Baums 2009: 140; Silverlock 2015: 231-232.

²⁰² The shaping of the junction between these footmarks and the basic characters varies between scribes, and can also be confused with the forms used to represent the *Cr* clusters. Refer to § 5.2 and § 5.4.2.1.

²⁰³ Salomon 2000: 82; Baums 2009: 140; Silverlock 2015: 232.

environment, the inclusion of apparent transitional forms such as *ḳ* and *g̣* suggests a more nuanced approach. It is possible that the scribes found it difficult to assess the true phonemic status and correct representation of the allophonic variations or sound changes occurring in the pronunciation of *k* and *g*, and the introduction of *ḳ* and *g̣* represents an attempt, albeit applied with some laxity, to resolve this situation by the RS and some other scribes. While the scribe was clearly aware of the lenition of the sounds, he was uncertain of the actual phonemic development and best way to represent this. In this context he may have chosen to retain the conservative forms of *k* and *g* in recognition of their historic status, while at the same time introducing the marked forms *ḳ* and *g̣* to indicate their weakened pronunciation. Although *y* is also used as a convenient grapheme with which to mark the lenition of these stops, this is less frequent than for the palatals *c* and *j* which are consistently recorded as *y* by all scribes, as discussed further in § 8.2.1.2.2. It is also possible that the degree of lenition varied with word position, and reflexes may have been intermixed in time through a combination of analogical spreading and scribal laxity. This is seen for example in the retention of *k* in some word initial forms following the *a*- prefix, suggesting a stronger pronunciation following the word initial vowel, and the lenition and deletion of suffixal *-ka* which may be related to the general weakening of word finals in Gāndhārī. In this context the extent of lenition may have been more limited in word medial positions, where the pronunciation could have been closer to that of a fricative as suggested. Whatever the original development or actual pronunciation of the recorded reflexes, the different patterns of usage seen between the BL and RS scribes suggest that these scholarly communities diverged to some extent in their graphic habits, and were possibly separated to some extent in location, time or dialect.

While most of the stops, including the voiceless stops, retain a contrastive phonemic status in modern day Torwali, they are usually weakened intervocalically, so that for example *g* is typically rendered as the velar fricative [ɣ], which also has a separate phonemic status in the language (Lunsford 2001: 9-26). Any intervocalic contrast between /*g*/ and /ɣ/ is thereby neutralized, creating difficulties in phonemic analysis of the language. The scribes in Gandhāra clearly faced similar problems, which were addressed by the use of *ḳ* and *g̣* for example. Although retention of weakened intervocalic stops in modern day Torwali is similar to the pattern observed in Gāndhārī, this does not necessarily indicate a strong historic connection between the two languages given the large time gap between the two records. This is discussed further § 8.2.6.3.

Aspirated voiceless velar stop:

Aś Ins: not recorded

BL, RS and BCS5 Scribes: $kh > \{h, (k), (kh), (g), (\emptyset)\} / V_V$

Aspirated voiced velar stop:

Aś Ins: $gh > \{h\} / V_V$

BL & BCS5 Scribes: $gh > \{h, (kh), (\emptyset)\} / V_V$

RS Scribe: $gh > \{g, (g), (\emptyset)\} / V_V$

Some of the examples given in the source texts are listed below:

Source		Sanskrit > Gāndhārī
Aś Ins	kh	not recorded
	gh	$gh > h$: <i>laghu-</i> > <i>lahu-</i>
BLS1-1	kh	$kh > h$: <i>sukham</i> > <i>suho</i> , <i>suduḥkhitaḥ</i> > <i>suduhidu</i> ,
	gh	$gh > h$: <i>laghu-</i> > <i>lahom</i> , <i>*jighatsābhiḥ</i> > <i>jihitsehi</i>
BLS1-3	gh	$gh > h$: <i>P oghatiṇṇo</i> > <i>ohatiṇo</i>
BLS2-2	kh	$kh > h$: <i>mukhareṇum</i> > <i>muha[r].n.</i> , <i>likhitakāni</i> > <i>lihidaga</i> , <i>sukhāsanena</i> > <i>suhasaneno</i>
BLS4-1	kh	$kh > h$: <i>*vīryamukhyāḥ</i> > <i>viriamuh[i]e</i>
	gh	$gh > h$: <i>oghānām</i> > <i>o[hāṇa]</i>
	gh	$gh > kh$: <i>anighaḥ</i> > <i>aṇikho</i> $gh > \emptyset$: <i>pratighaḥ</i> > <i>paḍia</i>
BLS9-1	kh	$kh > [h]$: <i>sukha-</i> > <i>su[ho]</i>
	gh	$gh > h$: <i>P appaṭigho</i> > <i>[a]paḍiho</i>
BCS5-1	kh	$kh > k$: <i>asamkhyeya</i> > <i>aṣakemaka[r]pa</i> $kh > h$: <i>pramukha-</i> > <i>pramuha-</i> , <i>sukha-</i> > <i>suha-</i> $kh > \emptyset$: <i>sukhe</i> > <i>suami</i>
	gh	$gh > h$: <i>laghūtthāna-</i> > <i>lahuṭṭhāna-</i> . BCS5-2: $gh > kh$: <i>āghāta-</i> > <i>akhada-</i>
RS-5	kh	$kh > h$: <i>sukhaye</i> > <i>suhae</i> $kh > g$: <i>nakhāḥ</i> > <i>ṇaga</i>
	gh	$gh > g$: <i>paridhovanīyaghaṭam</i> > <i>paradhoṇi-gada</i>

		<i>gh > g: dāghapālah > dagapale</i> <i>gh > Ø: dāgham > dae</i>
RS-14	<i>kh</i>	<i>kh > kh: akhilah > akhilo</i> <i>kh > g: akhilaiḥ > agilei, akhilena > agilena</i>
	<i>gh</i>	<i>gh > g: samghātī > -sagaḍi</i>
RS-19	<i>kh</i>	<i>kh > h: sukhite > suhide</i>
RS-20	<i>kh</i>	<i>kh > h: sukha- > suha-</i>
	<i>gh</i>	<i>gh > g: dīrgha-rātram > driga-ratro</i> (cluster with rhotic metathesis)

The typical development of *kh* involves de-occlusion as the aspirate *h*, which is consistently recorded by both the BL and RS scribes. De-occlusion of intervocalic *kh* appears to be common in other Gāndhārī documents such as the Dhp-G^K and ND, although the original form is sometimes retained (Salomon 2000: 82). While the BL and BC scribes have recorded a similar development for *gh > h*, the RS scribe has typically recorded the de-aspirated *g* or *ḡ* reflexes, seen in the RS-12, RS-14 and RS-20 manuscripts. The *h* reflex has not been observed for Sanskrit *gh* in these documents, and the impression is one of loss of aspiration leading to the merging of *-gh-* and *-g-* in this particular dialect (Silverlock 2015: 232). It is possible that the contrast *g : gh* was neutralized medially through deaspiration of the aspirate, which is reflected in spellings with *g*. This, in turn, allowed the possible use of *gh* to indicate a medial fricative, as discussed in relation to the unaspirated stops. The change appears to involve the sequential phonetic development of *-gh- > -g- > -ḡ-*²⁰⁴, which as suggested previously appears to have been used specifically to represent a weakened pronunciation, but one that was distinct from *y*. It is suggested that the single use of *kh* for *gh* by the BL scribe 4 is an erroneous retained historical spelling (Baums 2009: 143-144), while *Ø* is probably meant to represent *h*. Similar to the recording of the *y* reflex for the unaspirated stops, it is possible the use of *h* could also be a convenient available grapheme to record the weakened aspirated stops, while the use of *g* and *ḡ* by the RS scribe may represent an attempt to retain a contrastive reference to the original phoneme, distinct from *h*.

Based on the general hierarchy of articulatory strength presented by Gordon (2016: 153), aspirated stops are considered among the stronger elements, whereas a glottal gesture such as OIA *h* is among the weaker elements, similar to *y* on this scale. De-

²⁰⁴ The *ḡ* reflex is also recorded for *gh* in RS 12.

occlusion of a voiceless aspirated stop such as *kh* involves loss of the post-vocalic oral closure while the aspiration is retained. In the pronunciation of voiceless aspirated stops the voicing onset time (VOT) prior to voicing of the following vowel is lengthened, and the aspiration is realised with a more open laryngeal setting around the vocal cords (Ladefoged and Maddieson 1996: 66), which may be more similar to OIA visarga *ḥ* [h]. It has been suggested that the reflex *h* is pronounced as the voiced glottal fricative [ɦ] (Baums 2009: 143-144), consistent with OIA intervocalic *h* (§ 4.1), pointing towards an intermediate stage of voicing of the glottal gesture within the development. Assuming a staged process, then *kh* > *ḥ* > *h*, or *kh* > *gh* > *h* both seem possible. As observed above in relation to the unaspirated velar stop reflexes, the RS scribe appears to have been engaged with a more nuanced interpretation of the phonetic developments than the BL scribes, which is possibly reflected here in the divergent reflexes recorded for the voiced aspirate *gh*. In the pronunciation of voiced aspirated stops the voicing of the stop is maintained during the aspirated release, while the vocal folds appear to be held further apart, similar to voiceless stops, to facilitate the aspirated release of the stop (Ladefoged and Maddieson 1996: 69-70). The RS record suggests a perception of loss of the voiced aspiration as the initial stage of lenition in the process of *-gh-* > *-g-* > *-ḡ-*, but this interpretation is made without any great certainty.

8.2.1.2.2 Palatal stops

Unaspirated voiceless palatal stop:

Aś Ins: *c* > {*c*} / V _ V

BL, RS and BCS5 Scribes: *c* > {*y*, *Ø*, *c*, (*j*)} / V _ V

Unaspirated voiced palatal stop:

Aś Ins: *j* > {*j*, *y*, (*ch*)} / V _ V

BL, RS and BCS5 Scribes: *j* > {*y*, (*j*), (*c*), (*Ø*)} / V _ V

Some of the examples given in the source texts are listed below:

Source		Sanskrit > Gāndhārī
Aś Ins	<i>c</i>	not recorded
	<i>j</i>	<i>j</i> > <i>y</i> : <i>samaja-</i> > <i>samaya-</i> , <i>rājā-</i> > <i>raya-</i> <i>j</i> > <i>c</i> : <i>vraja-</i> > <i>vraca-</i>

BLS1-1	<i>c</i>	<i>c</i> > <i>y</i> : <i>acalam</i> > <i>ayalu</i> , <i>acaukṣam</i> > <i>ayokṣa</i> , <i>vācām</i> > <i>ba[y](*a)</i> , <i>aśuci</i> > <i>asuyi</i> , <i>vyarociṣam</i> > <i>viroyiṣu</i> <i>c</i> > <i>Ø</i> : <i>*paricaritvā</i> > <i>pariarita</i> , <i>sacet</i> > <i>sae</i>
	<i>j</i>	<i>j</i> > <i>y</i> : <i>pūjāyāḥ</i> > <i>puyae</i> , <i>bharadvājah</i> > <i>bharvayo</i> , <i>bhojanam</i> > <i>bhuyaṇo</i> , <i>rājagrhe</i> > <i>rayagha</i> <i>j</i> > <i>Ø</i> : <i>abhijānāmi</i> > <i>abhiaṇami</i> , <i>*tyajitvāna</i> > <i>caitaṇa</i> , <i>pravrajitam</i> > <i>parvaidu</i> , <i>BHS *prajānamānena</i> > <i>prabaeṣi</i>
BLS1-2	<i>c</i>	<i>c</i> > <i>y</i> : <i>-vacanam</i> > <i>-bayaṇo</i> , <i>-vacanatva-</i> > <i>bayaṇata</i> , <i>avaca</i> > <i>oya</i> <i>c</i> > <i>Ø</i> : <i>avaca</i> > <i>a[va]i</i> , <i>vicaya</i> > <i>vie</i>
BLS1-3	<i>c</i>	<i>c</i> > <i>y</i> : <i>P tacam</i> > <i>tvaya</i> / <i>śpaya</i>
	<i>j</i>	<i>j</i> > <i>Ø</i> : <i>P vanathajā</i> > <i>baṇahea[e]</i>
BLS2-1	<i>j</i>	<i>j</i> > <i>y</i> : <i>ājānīyam</i> > <i>ayaneja</i> , <i>pravrajitaḥ</i> > <i>pravayido</i> , <i>bhojanam</i> > <i>bhoyano</i>
BLS2-2	<i>j</i>	<i>j</i> > <i>j</i> : <i>rājā</i> > <i>raja</i> <i>j</i> > <i>y</i> : <i>rājā</i> > <i>raye</i> , <i>vrajati</i> > <i>vacadi</i> , <i>ājivika</i> > <i>āyiviga</i> , <i>mahārājah</i> > <i>maharaya</i>
BLS4-1	<i>c</i>	<i>c</i> > <i>y</i> : <i>gocaram</i> > <i>goyara-</i> , <i>vicikitsāḥ</i> > <i>viyigitsa-</i> , <i>śuciḥ</i> > <i>suyi</i> <i>c</i> > <i>Ø</i> : <i>brahmacaryiṣaṇā</i> > <i>bra(*maīyeṣana)</i>
	<i>j</i>	<i>j</i> > <i>y</i> : <i>*dravyajātikaḥ</i> > <i>dravayadia-</i> , <i>*parivrajanayā</i> > <i>parivrayaṇae</i> , <i>parivrajet</i> > <i>parivaye</i> , <i>ambujah</i> > <i>abuya</i> , <i>prayojayanti</i> > <i>prayoeti</i> , <i>bhojanam</i> > <i>bhoyaṇa</i> <i>j</i> > <i>Ø</i> : <i>niyojanam</i> > <i>ṇiyoaṇa</i>
BLS9-1	<i>c</i>	<i>c</i> > <i>y</i> : <i>gocarāya</i> > <i>go[yar.e]</i> , <i>no ce</i> > <i>ṇoya</i> , <i>vācābhilāpo</i> > <i>vayabh.lavo</i> , <i>sace</i> > <i>sayi</i> , <i>vicikiccham</i> > <i>(*vi)yigitsa</i>
	<i>j</i>	<i>j</i> > <i>y</i> : <i>bhajanti</i> > <i>bhayaṁti</i> , <i>bhujasmin</i> > <i>bhuyasi[m]</i> , <i>mahājanassa</i> > <i>mahayaṇaṣa</i>
BCS5-1	<i>c</i>	<i>c</i> > <i>y</i> : { <i>mocakah</i> / <i>mocayet</i> / <i>mocāya</i> } > <i>moyea</i> ; <i>upacaya-</i> > <i>uayea</i>
	<i>j</i>	<i>j</i> > <i>y</i> : <i>pravrajita-</i> > <i>parvayida-</i> <i>j</i> > <i>Ø</i> : <i>-bīja-</i> > <i>-bio/e-</i>
RS-5	<i>c</i>	<i>c</i> > <i>y</i> : <i>tvaca</i> > <i>tvaya</i>
	<i>j</i>	<i>j</i> > <i>j</i> : <i>ajānapadam</i> > <i>ajāṇavaḍa</i> <i>j</i> > <i>y</i> : <i>rajah</i> > <i>raya</i> <i>j</i> > <i>c</i> : <i>prajahatha</i> > <i>pracajaaṣa</i>
RS-12	<i>c</i>	<i>c</i> > <i>c</i> : <i>prasanna-cittam</i> > <i>(*pra)[s](*a)ṇa-cita</i> , <i>avi-cāram</i> > <i>avi-cara</i> <i>c</i> > <i>y</i> : <i>ārocayan</i> > <i>areyeṣi</i> , <i>sabrahma-cāribhiḥ</i> > <i>sabrama-yaraei</i> <i>c</i> > <i>Ø</i> : <i>ācakṣante</i> > <i>[a]ikṣati</i>
	<i>j</i>	<i>j</i> > <i>y</i> : <i>P samuttejesi</i> > <i>samuteyeṣi</i> , <i>avaskārabhājanataḥ</i> > <i>avakarabhayaṇaḍa</i>
RS-14	<i>c</i>	<i>c</i> > <i>Ø</i> : <i>sacet</i> > <i>sai</i>

	<i>j</i>	<i>j</i> > <i>y</i> : <i>pravrajitvā</i> > [p](r)(*a)[v](<i>a</i>)[yitva]
RS-19	<i>c</i>	<i>c</i> > <i>y</i> : <i>adhivacanam</i> > <i>aṣivayaṇa</i> ; <i>c</i> > <i>j</i> : P <i>nacirass-</i> > <i>ṇajiras(a)-</i> <i>c</i> > Ø: <i>sace</i> > <i>sae</i>
	<i>j</i>	<i>j</i> > <i>y</i> : <i>pravrajanti</i> > <i>parvayadī</i> , <i>prajānāmi</i> > <i>payañami</i>
RS-20	<i>j</i>	<i>j</i> > <i>y</i> : <i>prajānanti</i> > <i>payañati</i> ; <i>j</i> > <i>c</i> : <i>-jihvā</i> > <i>-cibha</i> (compound boundary)

The most widespread development of the intervocalic unaspirated palatal stops *c* and *j* involves lenition, expressed as the palatal approximant *y*, which is written consistently by all of the scribes. While reflexes involving elision (Ø) are also recorded in some manuscripts, this does not occur consistently within individual or across all texts. This often occurs in VCV sequences involving the vowel *i*, as in Skt. *abhijānāmi* > G *abhiaṇami* and *pravrajitam* > *parvaīdu* (BLS1-1) for example. Similar to the pattern observed for the velar stops, this suggests the scribes were aware of the phonetic similarity between *i* and *y*. Other cases, such as Skt. *niyojanam* > G *ṇiyoaṇa* (BLS4-1) where elision occurs between non-palatal vowels, may be due to an influence from the preceding *y*. Alternately, this could indicate the presence of a lightly audible expression of the *y* reflex as *ya-śruti* (see 8.2.1.2.1). which was not expressed by the scribes in a form of orthographic shorthand (Salomon 2008: 111). The original forms are also retained in some cases, typically following the verbal prefix *a-*²⁰⁵ or at compound boundaries²⁰⁶, but this is not a universal pattern. Other cases of the apparent retention of *j* are often reflexes of the *-ñj-* and *-jñ-* clusters (§§ 8.3.1.2, 8.3.2.1). The writing of *c* for *j* in *vraja-* > *vraca-* in the Aśokan inscriptions is an inverse spelling which supports the merger of *c* and *j*.

The overall impression is that intervocalic *c* and *j* were reduced to the palatal approximant *y*, which is mostly represented in the manuscripts by this advanced form, and it has been suggested that the velar and palatal unaspirated stops, *k*, *g*, *c* and *j*, had been reduced to a single pronunciation similar to *y* [j]²⁰⁷. The *y* reflex was used less frequently for the velar stops, however (§ 8.2.1.2.1), and based on the writing of modified stop reflexes by some of the scribes it was suggested reflexes such as *ga* may have been pronounced more like the voiced velar fricative [ɣ], representing a lesser extent of lenition for the velar stops. This raises questions about the general extent of lenition to the

²⁰⁵ Which represent both *a-* and *ā-* in the Kharoṣṭhī script.

²⁰⁶ Salomon 2008: 111; Silverlock 2015: 233.

²⁰⁷ Baums 2007: 143; Silverlock 2015: 233.

approximant *y* in Gāndhārī, and it is possible that *y* was used as a convenient grapheme to represent all of the weakened forms. The original extent of lenition may have also depended upon the word position, with subsequent intermixing of recorded reflexes due to analogical spreading or scribal laxity. Such a breakdown in the phonetic distinction and representation of four stops would have created some uncertainty amongst the scribes, who were no doubt aware of the loss of phonemic clarity resulting from the use of a single grapheme to represent at least five originally distinct sounds from the OIA languages. This would have led to the creation of many homonyms or even homographs depending upon the actual pronunciation, and may well have contributed to the use of historical or modified historical forms in certain cases (Silverlock 2015: 233-234). The use of modified forms for the velar stops by the RS scribe may also indicate an attempt to reconcile the actual phonetic and phonemic status with an appropriate underlying representation of these phonemes in their original word contexts, as discussed in § 8.2.1.2.1. Similar modified forms do not appear to have been introduced for the palatal reflexes, however, suggesting that these were likely reduced to the palatal approximant *y*, or alternately this form was used to create a distinction between the velar and palatal stops.

The Sanskrit aspirated voiceless palatal stop typically occurs in *-cch-* clusters in intervocalic positions (§ 8.3.1.1), while the voiced palatal stop *jh* is rare in intervocalic positions in Sanskrit, and even more so in Gāndhārī.

8.2.1.2.3 Retroflex stops

Unaspirated retroflex stops:

$t > \{ṭ, (ṭ)\} / V _ V$

$ḍ > ḍ / V _ V$ in limited examples

Aspirated retroflex stops:

$th > \{(dh)\} / V _ V$ in a single example

$dh > \{ḍh, (ḍ)\} / V _ V$ in limited examples

While OIA retroflex stops occur most commonly in clusters following the retroflex sibilant *ṣ* (§ 8.3.3.1, Appendix A1), they also occur as single consonants in intervocalic positions. Some of the examples given in the source texts are listed below:

Source	Sanskrit > Gāndhārī	Source	Sanskrit > Gāndhārī
BLS1-1	<i>t</i> > <i>ḍ</i> : <i>koṭiḥ</i> > <i>kuḍi</i> <i>varśakoṭīnām</i> > <i>barsakuḍiṇa</i>	BLS9-1	<i>ḍ</i> > <i>ḍ</i> : <i>saṃkrīḍanā</i> > <i>saṃkriḍaṇa</i> <i>ṭh</i> > <i>ḍh</i> : <i>śaṭhāḥ</i> > <i>śa[ḍha]</i>
BLS1-2	<i>ḍh</i> > <i>ḍh</i> : <i>muḍhaya</i> > <i>muḍhasa</i>	BCS5-1	<i>t</i> > <i>ḍ</i> : <i>koṭi</i> > <i>[koḍ]i-</i>
BLS2-2	<i>t</i> > <i>ḍ</i> : <i>kukkuṭā</i> > <i>kukuḍe</i> <i>BHS śīrśakaṭāhakam</i> > <i>ṣiṣakaḍahaga</i>	RS-5	<i>t</i> > <i>ḍ</i> : <i>kheṭaḥ</i> > <i>khade</i>
		RS-12	<i>t</i> > <i>ḍ</i> : <i>paridhovanīyaghaṭam</i> > <i>paradhoniḡaḍa</i>
BLS4-1	<i>t</i> > <i>ḍ</i> : <i>udghāṭanam</i> > <i>ugaḍaṇa</i> <i>bhutakoṭiḥ</i> > <i>bhudakoḍi</i> <i>ḍ</i> > <i>ḍ</i> : <i>ṣaḍindriyam</i> > <i>ṣaḍiidria</i> <i>pīḍām</i> > <i>piḍa</i>	RS-14	<i>t</i> > <i>ḍ</i> : <i>-saṃghāṭī</i> > <i>-sagaḍi</i>
		RS-20	<i>ḍ</i> > <i>ḍ</i> : <i>ayo-guḍaḥ</i> > <i>aya-uḍa</i> <i>biḍāla</i> > <i>biḷaḍa</i> (metathesis)

The unaspirated unvoiced stop *t* is typically voiced as *ḍ*, as in Skt. *koṭiḥ* > G *kuḍi* (BLS1-1) and *kheṭaḥ* > *khade* (RS-5) for example, although *ḍ* also arises as a reflex for dental *t* in many instances of the Skt. prefix *prati-* > G *paḍi-* (§ 8.2.1.2.4), where *prati* must have varied with metathesized *parti*, yielding retroflexion of *t*. Based on the single example of *śaṭhāḥ* > *śa[ḍha]* (BLS9-1) the unvoiced aspirated stop *ṭh* appears to become voiced as *ḍh*, while the latter is be retained in its original form, as in *muḍhaya* > *muḍhasa* (BLS1-2) for example. The presence of some Tocharian loanwords with similar pronunciation, along with sporadic confusions with the alveolar tap [ɾ] in the ND, point towards a weakened pronunciation of the retroflex stops as the retroflex tap [ɾ], which Baums (2009: 141) has suggested is supported by pronunciation in some of the modern Dardic languages.

8.2.1.2.4 Dental stops

Unaspirated voiceless dental stop:

Aś Ins: *t* > {*t*, *d*} / V _ V

BL Scribes: *t* > {*d*, (*d*), (*t*), (*Ø*)} / V _ V

RS & BCS5 Scribes: *t* > {*ḍ*, *ḍ*, *d/t*, (*ḍ*), (*Ø*), (*s*), (*ṣ*)} / V _ V

Unaspirated voiced dental stop:

Aś Ins: *d* > {*d*} / V _ V

BL Scribes: *d* > {*d*, (*Ø*), (*t*)} / V _ V

RS & BCS5 Scribes: *d* > {*ḍ*, *d/t*, (*d*), (*Ø*), (*ḍ*)} / V _ V

Some of the examples given in the source texts are listed below:

Source		Sanskrit > Gāndhārī
Aś Ins	<i>t</i>	Usually preserved, but some <i>t</i> > <i>d</i> : <i>hita-</i> > <i>hida-</i>
	<i>d</i>	Usually preserved (Mehendale 1948: 12).
BLS1-1	<i>t</i>	<i>t</i> > <i>t</i> : Limited examples: <i>ālapitaḥ</i> > <i>alavita</i> , <i>upasthitaḥ</i> > <i>upaṭhita</i> <i>t</i> > <i>d</i> : <i>agrataḥ</i> > <i>aghridu</i> , <i>anavatapte</i> > <i>aṇodatu</i> , <i>anubhūtam</i> > <i>aṇohodu</i> , <i>cyutaḥ</i> > <i>cudu</i> , <i>suduḥkhitaḥ</i> > <i>suduhidu</i> among many examples. <i>t</i> > <i>ḍ</i> : <i>pratibhānavān</i> > <i>paḍibhaṇavu</i> , <i>vivṛtāḥ</i> > <i>v[iv](*)ḍa</i>
	<i>d</i>	<i>d</i> > <i>d</i> : <i>idam</i> > <i>idu</i> , <i>udaram</i> > <i>udaru</i> , <i>nadītūram</i> > <i>ṇaditiru</i> , <i>yadā</i> > <i>yada</i> , <i>hṛdayam</i> > <i>hidaō</i> , among many examples.
BLS1-2	<i>t</i>	Typically, <i>t</i> > <i>d</i> : <i>ākhyātaḥ</i> > <i>akhade</i> , <i>anyataram</i> > <i>añadaro</i>
	<i>d</i>	<i>d</i> > <i>d</i> : <i>idam</i> > <i>ida</i>
BLS1-3	<i>t</i>	Typically, <i>t</i> > <i>d</i> : P <i>jahāti</i> > <i>jahadi</i> , P <i>vuccati</i> > <i>bucadi</i> <i>t</i> > <i>t</i> : P <i>vitatham idam</i> > <i>vita[ma]sea</i> <i>t</i> > <i>Ø</i> : P <i>uppatitam</i> > <i>upaīdo</i>
	<i>d</i>	<i>d</i> > <i>Ø</i> : <i>idam</i> > <i>ea</i>
BLS2-1	<i>t</i>	<i>t</i> > <i>d</i> : <i>āgacchati</i> > <i>agachadi</i> , <i>anyatarah</i> > <i>añadaro</i> , <i>utpādayati</i> > <i>upadayadi</i> , <i>jānāti</i> > <i>janadi</i> , <i>yataḥ</i> > <i>yado</i> among many examples.
BLS2-2	<i>t</i>	<i>t</i> > <i>d</i> : <i>antarhitaḥ</i> > <i>atarahido</i> , <i>ghītaḥ</i> > <i>grahad[o]</i> , <i>caturthaḥ</i> > <i>caduṭho</i> , <i>parataḥ</i> > <i>parado</i> , <i>vahati</i> > <i>bahadi</i> among many examples.
BLS4-1	<i>t</i>	<i>t</i> > <i>d</i> : <i>anupahata-</i> > <i>aṇuahada-</i> , <i>bhūtakoṭiḥ</i> > <i>bhudakoḍi-</i> among others. <i>t</i> > <i>ḍ</i> : <i>pratighaḥ</i> > <i>paḍia</i> , <i>prati-</i> > <i>paḍi-</i> in several examples.
	<i>d</i>	<i>d</i> > <i>d</i> : <i>*śāntivāḍaiḥ</i> > <i>śatipado</i> , <i>sadā</i> > <i>sada</i> , <i>*samudaya-</i> > <i>samudaga-</i> , <i>[pa]do</i> > <i>padam</i> among others.
BLS9-1	<i>t</i>	Usually retained, <i>t</i> > <i>t</i> : <i>icchati</i> > <i>achati</i> , <i>upātivatto</i> > <i>uvativotto</i> , <i>kuto</i> > <i>[k]uto</i> , <i>dutiyena</i> > <i>dutieṇa</i> , <i>vāto</i> > <i>vato</i> among many examples. <i>t</i> > <i>d</i> : <i>cataso</i> > <i>cadaso</i> , <i>-ratassa</i> > <i>-radasa</i> , <i>sañjāta</i> > <i>saṃjada</i> <i>t</i> > <i>ḍ</i> : <i>prati-</i> > <i>paḍi-</i>
	<i>d</i>	<i>d</i> > <i>d</i> : <i>sudurlabhā</i> > <i>sudalabho</i> <i>d</i> > <i>t</i> : <i>idāni</i> > <i>taṇi</i> , <i>padumī</i> > <i>patumam</i> , <i>muditañ</i> > <i>mutita</i>
BCS5-1	<i>t</i>	<i>t</i> > <i>d</i> : <i>aṇāgata</i> > <i>aṇagada</i> , <i>atīta-</i> > <i>adida-</i> , <i>itara-</i> > <i>idara-</i> , <i>catur</i> > <i>cadura</i> , <i>cintita-</i> > <i>citida-</i> , <i>jāti-</i> > <i>jadi-</i> , <i>viṃśati</i> > <i>viṣṭadi</i> among others. <i>t</i> > <i>ḍ</i> , <i>d</i> , <i>Ø</i> : <i>prati-</i> > <i>paḍi</i> / <i>pradi</i> / <i>prai</i> <i>t</i> > <i>s</i> : <i>svabhāvatā</i> > <i>śpabhavasa</i> ; <i>t</i> > <i>Ø</i> : <i>aciniteṇa</i> > <i>acitieṇa</i>
	<i>d</i>	<i>d</i> > <i>d</i> : <i>yadi</i> > <i>yadi</i>
RS-5	<i>t</i>	Typically, <i>t</i> > <i>ḍ</i> : <i>katarah</i> > <i>kaḍara</i> , among many cases.

		<i>t > d: bhavati > bhodi</i> <i>t > ḍ: pratikūla > paḍikula</i> <i>t > s: jānataḥ > jaṇasa</i> <i>t > ṣ: paśyataḥ > paśaṣa</i> <i>t > Ø: anyatarah > aṇeare</i>
	<i>d</i>	Typically, <i>d > ḍ: pādatalāt > paḍatala</i> , among many cases. Some exceptions such as <i>d > d: paridāmyati > paridame[ḍ]i</i>
RS-12	<i>t</i>	Typically, <i>t > ḍ: sṭhitā > aṭhide</i> , <i>adhigataḥ > aṣigaḍe</i> , <i>prīteḥ > piḍia</i> , BHS <i>pratisamviditāḥ > paḍisobiḍa</i> , <i>rahogataḥ > rahogaḍa</i> , <i>samāhitam > samahide</i> , <i>prītisukham > piḍisuho</i> , among many examples. Also, <i>t > ḍ: prati- > paḍi-</i>
	<i>d</i>	Typically, <i>d > ḍ: idam > iḍam</i> , <i>pradeśam > praḍeṣi</i> , <i>yāvad eva > yavaḍe</i>
RS-14	<i>t</i>	<i>t > ḍ: ācaritam > -(a)carido</i> , <i>kṛtam > kiḍe</i> , <i>icchate > chaḍe</i> , <i>vicintitam > vicitido</i> , among others.
	<i>d</i>	<i>d > ḍ: P adāsi > aḍa[s]i</i> , <i>P tādīnā > taḍīṇa</i> , regular without exception.
RS-19	<i>t</i>	<i>t > ḍ: -pūti > -puḍi</i> , <i>P carati > caraḍi</i> , <i>duḥkhite > duhide</i> , <i>avavāditah > ovaḍido</i> , among others. <i>t > Ø: anyatarah > aṇeare</i>
	<i>d</i>	<i>d > ḍ: avavādena > ovaḍeṇa</i> , <i>avavāditah > ovaḍido</i> <i>d > Ø: P niyyāditā > ṇiāiḍa</i> , <i>samsadanam > saṣiaṇa</i>
RS-20	<i>t</i>	Typically, <i>t > ḍ: abhiratāḥ > aviraḍa</i> , <i>etat > eḍaḍ</i> , among many others. <i>t > t / d: utpadyate > apacadi / apacati</i> <i>t > ḍ: pratipad > paḍipad</i> , similar to other scribal examples.
	<i>d</i>	Typically, <i>d > ḍ: idam > iḍam</i> , <i>utpādāya > upaḍae</i> , among many others. <i>t > ḍ: āḍīptāḥ > aḍita</i> , <i>paridāhaḥ > paraḍae / paraḍao</i> <i>d > d: sukha-duḥkham > suha-dukho</i> , typical at compound boundaries.

The most commonly occurring reflexes express the voicing of *t > d* and the preservation of *d*, which together suggest the merger or loss of distinction between *t* and *d*. These are both typically recorded with the diacritic marked *ḍ* by the RS scribe²⁰⁸. Other reflexes are seen on a limited basis, including the retention of historic forms at compound boundaries, as in Skt. *sa-devakasya > G sadevayesa* for example (Silverlock 2015: 235-6), the irregular use of *s* and *ṣ* for *t* by the RS scribe, which may be due to orthographic confusion with the aspirated stop reflexes (Glass 2007: 116), or a fricative pronunciation, and recording of the Sanskrit verbal prefix *prati-* as *paḍi-* on a regular basis across all of the Gāndhārī manuscripts.

²⁰⁸ The use of this diacritic form has also been observed in manuscripts of the New Collection, which have not been published in a comprehensive form to date, and for this reason are not included in this thesis (§3.3).

While the voicing of *t* and merger with *d* is the most frequent Gāndhārī development (e.g., Allon 2001: 82), original *t* is retained on a more regular basis in the Khvs-G, written by the BL scribe 9, and this inconsistency is assumed to reflect the loss of distinction between *t* and *d*, along with loose orthographic practices (Salomon 2000: 84). The regular use of G *paḍi*- for OIA *prati*- appears to represent a form borrowed from an earlier MIA dialect which was adopted into the language, especially as *pr*- is regularly retained in Gāndhārī (Silverlock 2015: 235-236), as for instance in alternate reflexes such as *prati*- (§ 8.3.1.4.4). It is suggested that the *ḍ* developed under the influence of preceding *r*, as in *t > ṭ* (in Pali) *> ḍ*, while *r* was lost and *ḍ* was retained in subsequent MIA developments. The spread of retroflexion to dental stops from preceding *r* is not recognised in the regular Sanskrit sandhi rules (Appendix A1), and this development appears to represent a form of spontaneous retroflexion (Marino 2017: 122), which is also a common MIA phenomenon. Other similar examples which are consistent with the regular sandhi rules include Skt. *apratipudgalah* *>* G *apradipughalu*, *hṛdayam* *>* *hidao* (Salomon 2008: 113-114), and *kṛtam* *>* *kide* (Salomon 2000: 339) for example, supporting the interpretation that *paḍi* is an irregular language development which was somehow adopted on a widespread basis.

The regular use of the reflex *ḍ* by the RS scribe is interpreted as indicating a weakening in pronunciation, possibly as the voiced dental fricative [ḍ]²⁰⁹. As noted in § 8.2.1.2.1 the RS scribe was possibly more engaged with a finer interpretation of the phonetic and phonemic distinctions than the BL scribes, which may account for the use of the diacritic marking to this character. At the same time the original phonemic form of *t/d* is retained by all scribes, although to varying extents, suggesting that the single dental stops retained a distinct phonemic status in intervocalic positions. This compares with the apparent confusion created by the regular use of the grapheme *y* to mark the weakened forms of both the velar and especially the palatal intervocalic stops (§§ 8.2.1.2.1, 8.2.1.2.2, 8.2.1.2.6).

Aspirated voiceless dental stop:

Aś Ins: not recorded

BL, RS and BCS5 Scribes: *th* *>* {*s*, *ṣ*, (*th*), (*ḍh*), (*ḍh*), (*ḍ*)} / V _ V

²⁰⁹ Salomon 2008: 339; Baums 2009: 141-142; Silverlock 2015: 234; Marino 2017: 121.

Aspirated voiced dental stop:

Aś Ins: *dh* > {*dh*} / V _ V

BL, RS and BCS5 Scribes: *dh* > {*s*, *ṣ*, (*dh*)} / V _ V

Some of the examples given in the source texts are listed below:

Source		Sanskrit > Gāndhārī
Aś Ins	<i>th</i>	none recorded
	<i>dh</i>	none recorded
BLS1-1	<i>th</i>	<i>th</i> > <i>s</i> : <i>gāthākārāṇām</i> > <i>ghasagharāṇa</i> , <i>tapiṣyatha</i> > <i>taviśasa</i> , <i>tathā</i> > <i>tasa</i> , <i>mokṣyatha</i> > <i>mukṣasa</i> , <i>pr̥thivim</i> > <i>prasavi</i>
	<i>dh</i>	<i>dh</i> > <i>s</i> : <i>anadhikāni</i> > <i>aṇasia</i> , <i>madhura</i> > <i>masura</i>
BLS1-2	<i>th</i>	<i>th</i> > <i>s</i> : <i>yathā</i> > [<i>yasa</i>], <i>athāparam</i> > <i>hasavaro</i> , BHS <i>sayyathāpi</i> > <i>sayasavi</i>
	<i>dh</i>	<i>dh</i> > <i>s</i> : <i>adhikaraṇam</i> > <i>asiraṇa</i> , <i>nirodha-</i> > <i>ṇirusa-</i> , <i>damatha-</i> > [<i>da</i>] <i>masa-</i>
BLS1-3	<i>th</i>	<i>th</i> > <i>s</i> : P <i>-kathamkatho</i> > <i>-kasakasa</i> , P <i>vitatha</i> > <i>vitasa-</i> <i>th</i> > <i>h</i> : P <i>vanathajā</i> > <i>baṇahea[e]</i>
BLS2-1	<i>th</i>	<i>th</i> > <i>ṣ</i> : <i>yathā-</i> > <i>yaṣa-</i> ; also, <i>th</i> > <i>th</i> : <i>gāthā</i> > <i>gatha</i>
	<i>dh</i>	<i>dh</i> > <i>ṣ</i> / <i>s</i> : <i>pranidhiḥ</i> > <i>praniṣi</i> , <i>bodhisattvaḥ</i> > <i>boṣisatvo</i> / <i>bosisatva</i>
BLS2-2	<i>th</i>	<i>th</i> > <i>ṣ</i> : <i>yathā-</i> > <i>yaṣa-</i> , <i>yathopamānam</i> > <i>yaṣayupamano</i> , <i>*dākṣiṇapathānām</i> > <i>dakṣiṇavaṣano</i>
	<i>dh</i>	<i>dh</i> > <i>ṣ</i> / <i>s</i> : <i>pratyekabodhim</i> > <i>pracageboṣ(*i)</i> , <i>bodhisattva-</i> > <i>bosisatva-</i>
BLS4-1	<i>th</i>	<i>th</i> > <i>s</i> : <i>atha</i> > <i>asa</i> , <i>karmapathānām</i> > [<i>ka</i>] <i>ma[pa]saṇa</i> , <i>yathā</i> > <i>yasa</i> . <i>th</i> > <i>ḍh</i> : <i>prathamam</i> > <i>paḍhamo</i>
	<i>dh</i>	<i>dh</i> > <i>s</i> : <i>adhīnaḥ</i> > <i>asiṇo</i> , <i>adhimātram</i> > <i>asimatro</i> , <i>bodhiḥ</i> > <i>bosi</i> , <i>bodhisattva-</i> > <i>bosisatva-</i> , <i>medhāvī</i> > <i>misavi</i> , <i>samādhīḥ</i> > <i>samasi</i> , <i>sumedhaḥ</i> > <i>sumesu</i> ,
BLS9-1	<i>th</i>	<i>th</i> > <i>ṣ</i> : <i>yathā</i> > <i>yaṣa</i> / [<i>yasa</i> one time], <i>yūthāni</i> > <i>yusaṇi</i>
	<i>dh</i>	<i>dh</i> > <i>ṣ</i> : <i>kṣudham</i> > <i>kṣuṣa</i> , <i>nidhāya</i> > <i>ṇiṣae</i> , <i>madhurā</i> > <i>maṣora</i>
BCS5-1	<i>th</i>	<i>th</i> > <i>ṣ</i> : <i>yathā</i> > <i>yaṣa</i> , <i>katham</i> > <i>kaṣa</i> <i>th</i> > <i>ḍh</i> : <i>prathama-</i> > <i>paḍhama-</i>
	<i>dh</i>	<i>dh</i> > <i>ṣ</i> : <i>adhi-</i> > <i>aṣi-</i> , <i>bodhimaṇḍa-</i> > <i>boṣimada-</i> , <i>bodhisattva-</i> > <i>boṣisatva-</i> , <i>asādhāraṇa-</i> > <i>asaṣaraṇa-</i> . But <i>sudha</i> > <i>sudha</i> uncertain?
RS-5	<i>th</i>	<i>th</i> > <i>ṣ</i> / <i>s</i> : <i>yathā</i> > <i>yaṣa</i> , but also <i>atha</i> > <i>asa</i>
	<i>dh</i>	<i>dh</i> > <i>ṣ</i> / <i>s</i> : <i>samādhīḥ</i> > <i>ṣamaṣi</i> , <i>adhaḥ</i> > <i>asa</i>

		<i>dh > dh: anudharmaḥ > aṇudharma</i>
RS-12	<i>th</i>	<i>th > s / ṣ: atha > asa / aṣa, yathā- > yaṣa-, -kathā- > -kaṣa- th > ḍ: prathamam > praḍama</i>
	<i>dh</i>	<i>dh > ṣ: adhigata- > aṣigaḍa-, P sādhu > sasū / sasa</i> After preverb or at compound boundaries, <i>dh > dh: pari-dhovanīya-ghaṭam > para-dhoṇigaḍa, mānuṣya-dharmāt > maṇuṣa-dharma</i>
RS-14	<i>th</i>	<i>th > s / ṣ: yathā > yasa / yaṣa</i>
	<i>dh</i>	<i>dh > s: madhuraṁ > masuro</i>
RS-19	<i>th</i>	<i>th > s: atha > asa</i>
	<i>dh</i>	<i>dh > s / ṣ: praṇidhāya > paṇisae; adhivacanam > aṣivayaṇa</i>
RS-20	<i>th</i>	<i>th > s / ṣ: yathā > yaṣa; atha > asa, BHS sayyathāpi > sayasavi</i>
	<i>dh</i>	<i>dh > s / ṣ: vividham > viviso, duḥkha-ñirodhaḥ > dukha-ñiroṣo</i>

The most commonly recorded reflexes of both *-th-* and *-dh-* are the sibilants *s* and *ṣ*, with the addition of an underscore diacritic in the latter case. The use of these alternate reflexes appears to vary randomly between and within texts, with the BL scribe 1 using ordinary *s* almost exclusively, while graphic alternation between the ordinary and diacritic forms is common in the work of other BL scribes and the BC and RS scribes²¹⁰. Less common, but recorded by several scribes, is the development of *th > ḍh / ḍ* in Skt. *prathama-* > G *paḍhama-* / *praḍama-* (e.g., BLS4-1, RS-12), which is assumed to be influenced by the preceding *r*²¹¹, while *dh* is recorded in several instances at morpheme or compound boundaries²¹². Irregular reflexes such as the de-occlusion of *th > h* and the retention of *th* (Lenz 2003: 42, 130) are only represented in single cases.

The recorded reflexes suggest lenition of the stops to form a fricative in a process of spirantisation (Kirchner 2004: 313-314). It is suggested that this was most likely pronounced as the voiced alveolar fricative [z]²¹³, while an intermediate stage as the voiced dental fricative [ð] has also been suggested²¹⁴. From the review of extensive language surveys, Kirchner (2004: 316-317) has observed that unaffricated stops never spirantise synchronically to a sibilant or labiodental fricative, suggesting a staged process involving early affrication, which may have also corresponded with the loss of aspiration. The homorganic [ts] affricate is recorded in Torwali (Lunsford 2001: 11, 22), and further

²¹⁰ Allon 2001: 83; Lenz 2003: 42; Salomon 2008: 114, 339.

²¹¹ Baums 2009: 145; Silverlock 2015: 237; Schlosser 2016: 75.

²¹² Glass 2007: 117; Silverlock 2015: 237-238.

²¹³ Salomon 2000: 85; Allon 2001: 83; Lenz 2003: 43; Baums 2009: 145

²¹⁴ Brough 1962: § 43, 94; Silverlock 2015: 237-238.

weakening of such a sound towards the unvoiced dental fricative *s* [θ], or its voiced development as [ð] in Gāndhārī, would not be unexpected. Since this sound is associated with the lenition of *t* and *d*, which developed reflexes distinct from those of *th* and *dh*, it has been suggested that all four sounds would have merged as [z] if their lenition was synchronic (Baums 2007: 144-145). The different reflexes therefore suggest that the aspirates *th* and *dh* may have weakened directly to [z], or else weakened prior to the non-aspirates, which only developed as far as [ð] (Silverlock 2015: 238). The distinction between *s* and *ś* may be one of diachronic sequence, with *ś* possibly representing an intermediate sound such as [ð], while *s* represents the advanced development as [z] (Silverlock 2015: 237). Within this context it is possible that *ś* was introduced originally to distinguish the developing allophones of phonemes *th* and *dh* from the existing sibilant phoneme *s*, as a way of confirming their distinct phonemic status. This would have avoided the confusions arising from the use of a single grapheme to represent several phonemes, similar to the interpretation of *g* discussed in § 8.2.1.2.1. This is supported by the most frequent usage occurring in the hand of the RS scribe, who as mentioned earlier (e.g., § 8.2.1.2.1) appears to have been more engaged with these finer distinctions than the BL scribes. It is possible that as the pronunciation merged more closely with the fricatives, whether dental or alveolar, the distinction between the two forms gradually merged in the hands of all scribes through a process of analogical spreading and scribal laxity, which may have also been associated with a leveling of the sounds at a later stage in the language development (Salomon 2000: 85). A similar pattern is also seen in the representation of the fricative sibilants (§ 8.2.3).

8.2.1.2.5 Labial stops

Unaspirated voiceless labial stop:

Aś Ins: $p > \{p, v\} / V _ V$

BL, RS and BCS5 Scribes: $p > \{v, \emptyset, p, (b)\} / V _ V$

Unaspirated voiced labial stop:

Aś Ins: not recorded

BL, RS and BCS5 Scribes: $b > \{v, \emptyset, (b)\} / V _ V$

Some of the examples given in the source texts are listed below:

Source		Sanskrit > Gāndhārī
Aś Ins	<i>p</i>	<i>p</i> > <i>p</i> : <i>apatya-</i> > <i>apaca-</i> (Woolner 1924, Part II. Glossary: 63) <i>p</i> > <i>v</i> : <i>pra-√āp</i> > <i>pa-√āv(a)</i> (Mehendale 1948: 13); <i>apa-</i> > <i>ava-</i> (Hultsch 1925: lxxxv)
	<i>b</i>	not recorded
BLS1-1	<i>p</i>	Typically, <i>p</i> > <i>v</i> : <i>anulepanena</i> > <i>aṇol[e]vaṇena</i> , <i>ālapitaḥ</i> > <i>alavita</i> , <i>kuṇapam</i> > <i>kuṇavu</i> , <i>pāpikām</i> > <i>pavia</i> , <i>vinipātam</i> > <i>vinivadu</i> , <i>upadrutaḥ</i> > <i>vadhrodu</i> ; <i>p</i> > <i>p</i> : <i>upasthitaḥ</i> > <i>upaḥhita</i> , <i>*kṣipā</i> > <i>kṣip[u]</i> <i>p</i> > <i>Ø</i> : <i>antaḥpure</i> > <i>ateuru</i> , <i>abhirūpaḥ</i> > <i>abhiroo</i> , <i>upananaddhaḥ</i> > <i>oaṇadho</i> , <i>rūpāṇi</i> > <i>rua</i> . Elision is often associated with the <i>u</i> vowel. <i>p</i> > <i>b</i> : <i>stūpam</i> > <i>thubu</i> – possibly another dialect?
	<i>b</i>	<i>b</i> > <i>v</i> : BHS <i>*nirbandhaye</i> > <i>ṇivadho</i> ; <i>b</i> > <i>Ø</i> : <i>vibudhya</i> > <i>viojita</i>
BLS1-2	<i>p</i>	Typically, <i>p</i> > <i>v</i> : <i>āpadyate</i> > <i>avajadi</i> , <i>pāpakāḥ</i> > <i>pavea</i> , <i>rūpa-</i> > <i>-rova-</i> , BHS <i>sayyathāpi</i> > <i>sayasavi</i> , <i>athāparam</i> > <i>hasavaro</i> <i>p</i> > <i>p</i> : BHS <i>vipūyakam</i> > <i>v[i]puao</i> , P <i>anāthapiṇḍikassa</i> > <i>aṇasapiḍiasa</i> <i>p</i> > <i>Ø</i> : P <i>upasaṅkami</i> > <i>uasakrami</i> , <i>upekṣā</i> > <i>uekṣa</i> , <i>rūpāṇi</i> > <i>ruaṇa</i>
BLS2-1	<i>p</i>	<i>p</i> > <i>v</i> : <i>āpannaḥ</i> > <i>avarnage</i> , <i>nidhāpitaḥ</i> > <i>nidhavido</i> , <i>vipākāḥ</i> > <i>vivago</i> <i>p</i> > <i>p</i> : <i>upakaraṇam</i> > <i>upagarano</i> , <i>upamānam</i> > <i>upamano</i>
BLS2-2	<i>p</i>	<i>p</i> > <i>v</i> : <i>aparaḥ</i> > <i>avare</i> , <i>upagataḥ</i> > <i>uvagado</i> , <i>gopālikā</i> > <i>govalige</i>
	<i>b</i>	<i>b</i> > <i>v</i> : <i>sabalaḥ</i> > <i>-savalo</i> <i>b</i> > <i>Ø</i> : P <i>pāpunissaṃ</i> > <i>praṇi[śa]</i>
BLS4-1	<i>p</i>	<i>p</i> > <i>v</i> : <i>aparāddhaḥ</i> > <i>avaradho</i> , <i>aparo</i> > <i>avaro</i> , <i>janapada-</i> > <i>janavada-</i> , <i>nipakānām</i> > <i>ṇivaana</i> , <i>pratipadam</i> > <i>paḍivada</i> , <i>pāpakāḥ</i> > <i>pavaga</i> , <i>pāpam</i> > <i>pavo</i> , <i>vyāpāda-</i> > <i>vavada</i> <i>p</i> > <i>Ø</i> : <i>anupagatāḥ</i> > <i>aṇuagada</i> , <i>anupādānaḥ</i> > <i>aṇuadana</i> , <i>*upapādāyatane</i> > <i>uavadayadane</i>
	<i>b</i>	<i>b</i> > <i>v</i> : <i>abandhanaḥ</i> > <i>avadhaṇo</i>
BLS9-1	<i>p</i>	Typically, <i>p</i> > <i>v</i> : P: <i>upakkilese</i> > <i>uvakileśa</i> , <i>upātivatto</i> > <i>uvativuto</i> , <i>upekham</i> > <i>uveka</i> , <i>nipakam</i> > <i>ṇivago</i> , <i>pipāsam</i> > <i>pivaṣa</i> <i>p</i> > <i>Ø</i> : <i>nillolupo</i> > <i>ṇiloluo</i>
	<i>b</i>	<i>b</i> > <i>b</i> : <i>abaddhaḥ</i> > <i>abadho</i> <i>b</i> > <i>v</i> : <i>pratibaddha-</i> > <i>paḍi[v]adha-</i> <i>grhivyaṃjanāni</i> > <i>gihibamdhanaṇi</i> (uncertain).
BCS5-1	<i>p</i>	Typically, <i>p</i> > <i>v</i> : <i>anupādāna</i> > <i>aṇu[va]daṇa</i> , <i>kapālaka</i> > <i>kavalaa</i> , <i>pāpa</i> > <i>pava</i> , <i>apacaya</i> > <i>avayeasa</i> , <i>apanaya</i> > <i>avaṇao</i> <i>p</i> > <i>Ø</i> : <i>upa-</i> > <i>ua-</i>
BCS5-2	<i>b</i>	<i>b</i> > <i>v</i> : <i>sabala-</i> > <i>savala-</i>

RS-5	<i>p</i>	Typically, <i>p</i> > <i>v</i> : <i>ajanapadam</i> > <i>ajāṇavaḍa</i> <i>p</i> > \emptyset : <i>rūpam</i> > [r]uo
RS-12	<i>p</i>	Typically, <i>p</i> > <i>v</i> : <i>-saṃlāpam</i> > <i>-śalava</i> , <i>upa-</i> > <i>uva-</i> , <i>kulapari-</i> > <i>kulavara-</i> , <i>pāpakaiḥ</i> > <i>pavaehi</i> , <i>yāpanīyam</i> > <i>yava</i> (*nio) <i>p</i> > <i>p</i> : <i>kula-putrāḥ</i> > <i>kulaputra</i> , <i>dāghapālāḥ</i> > <i>dagapale</i> <i>p</i> > \emptyset : <i>upa-</i> > <i>ua-</i> , <i>kulaputraḥ</i> > <i>kulaü</i> (*tra), <i>rupa-</i> > <i>rua</i>
	<i>b</i>	<i>b</i> > <i>v</i> : BHS <i>anubandhitvā</i> > <i>aṇuvadhita</i>
RS-14	<i>p</i>	<i>p</i> > <i>v</i> : <i>kadāpi</i> > (*kaḍa)vi, <i>kāsyapaḥ</i> > [ka]śave, <i>praṇipatyā</i> > (*pra)ṇavaca, <i>vipākāḥ</i> > <i>vivace</i>
RS-19	<i>p</i>	Typically, <i>p</i> > <i>v</i> : P <i>āpajjati</i> > <i>avajaḍi</i> , <i>aparam</i> > <i>avara</i> , <i>kimapi</i> > <i>kicavi</i> , <i>nāparam</i> > <i>ṇavara</i> <i>p</i> > \emptyset : <i>gopālakaḥ</i> > <i>goalao</i> , BHS <i>anupādāya</i> > <i>aṇuḍae</i>
RS-20	<i>p</i>	<i>p</i> > <i>v</i> : <i>pratipad</i> > <i>paḍivaḍa</i> <i>p</i> > <i>p</i> : <i>jāti-paridāham</i> > <i>jaḍi-paraḍaa</i> , but mostly compound boundaries. <i>p</i> > \emptyset : <i>upa-</i> > <i>ua-</i>

The most commonly recorded reflex for both *-p-* and *-b-* is the labial approximant *v*, as in Skt. *ālapitaḥ* > G *alavita* (BLS1-1) and *sabala-* > *savala-* (BCS5-1) for example, while elision is also recorded in a significant number of cases, such as in *rūpāṇi* > *rua* (BLS1-1) and *gopālakaḥ* > *goalao* (RS-19). Similarly, while the common prefix *upa-* is sometimes rendered as *uva-*, as in *upagataḥ* > *uvagado* (BLS2-2) for example, it is often rendered as *ua-* or *oa-*, besides being retained as *upa-*, as in *upakaraṇam* > *upagarano* (BLS2-1). The alternation between *v* and \emptyset is more or less typical in Gāndhārī, and appears to reflect a merger of *p* and *b* with *v*²¹⁵. In most cases elision appears to follow an orthographic rule in Gāndhārī that *v* is unwritten when one of the flanking vowels is *u* / *o*, in what is a graphic shorthand rather than true elision²¹⁶. Use of the reflex *v* is assumed to represent the phonetic spelling of a weakened pronunciation which is close to the voiced labial approximant [v]²¹⁷. Other reflexes include the retention of *p* in a number of cases which appear to be historicisms used at compound boundaries²¹⁸, while the voicing recorded in Skt. *stūpa-* > *thubu-* may reflect borrowings (Salomon 2008: 115).

Similar to the intervocalic palatal stops which are typically represented by the phoneme *y* (§ 8.2.1.2.2), it is assumed that the weakened bilabial stops have a similar pronunciation the existing phoneme *v* which has been used to represent them. While this

²¹⁵ Salomon 2008: 115; Silverlock 2015: 239; Schlosser 2016: 75-76.

²¹⁶ Baums 2009: 142; Silverlock 2015: 238.

²¹⁷ Baums 2009: 143; Silverlock 2015: 238.

²¹⁸ Silverlock 2015: 238; Marino 2017: 123.

suggests a full merger between these phonemes in intervocalic positions, it remains to be understood whether the original *p* and *b* retained a distinct phonemic status in medial positions in Gāndhārī. This assessment is complicated by the use of the single forms of these graphemes to represent geminates (§ 8.3.1.1), as well as the reflexes of clusters such as *-ṃT-* for example (§ 8.3.2.1), so that along with approximating the weakened sound of the stops, the use of *v* was possibly a convenient way to mark this distinction. Evidence from the other stop series supports a merger of voiced and voiceless stops in favour of voicing, however, and *b* may represent a stop-stage similar to that which occurs for the other stops, while *v* is the final fricative stage.

No reflexes of the OIA aspirated voiceless labial stop *ph* in intervocalic positions are recorded in the sources, although this phoneme occurs in some reflexes of the OIA *sp* cluster (§ 8.3.3.1).

Aspirated voiced labial stop:

Aś Ins: *bh* > {*bh*, *h*} / V _ V

BL, RS and BCS5 Scribes: *bh* > {*bh*, *h*, *v*, (Ø), (*vh*), (*y*)} / V _ V

Some of the examples given in the source texts are listed below:

Source		Sanskrit > Gāndhārī
Aś Ins	<i>bh</i>	<i>bh</i> > <i>h</i> : <i>-bhiḥ</i> > <i>-hi</i> (Mehendale 1948: 13)
BLS1-1	<i>bh</i>	<i>bh</i> > <i>h</i> : <i>anubhūtam</i> > <i>aṇohodu</i> , <i>devabhūtaḥ</i> > <i>dev[ah]odu</i> , <i>vaśībhūtaḥ</i> > <i>baīhodu</i> , <i>durbhikṣe</i> > <i>dhrohikṣu</i> <i>bh</i> > <i>bh</i> : <i>abhi-</i> > <i>abhi-</i> <i>bh</i> > Ø: <i>aśubham</i> > <i>aśua</i>
BLS1-2	<i>bh</i>	<i>bh</i> > <i>bh</i> : <i>abhi-</i> > <i>abhi-</i>
BLS4-1	<i>bh</i>	<i>bh</i> > <i>h</i> : <i>lobha-</i> > <i>loha</i> , <i>vimuktibhiḥ</i> > <i>vimutihi</i> , <i>samāpattibhiḥ</i> > <i>samavatihi</i> <i>bh</i> > <i>bh</i> : <i>abhi-</i> > <i>abhi-</i> <i>bh</i> > <i>v</i> : <i>abhidhyāyati</i> > <i>avijayadi</i>
BLS9-1	<i>bh</i>	<i>bh</i> > <i>bh</i> : <i>vācābhilāpo</i> > <i>vayabhilavo</i> <i>bh</i> > <i>h</i> : <i>*abhiyācaye</i> > <i>ahiacae</i>
BCS5-1	<i>bh</i>	<i>bh</i> > <i>bh</i> : <i>paribhāṣā-</i> > <i>paribhaṣa-</i> , <i>paribhuj-</i> > <i>paribhuj-</i> , <i>lābha-</i> > <i>labha-</i> <i>bh</i> > <i>vh</i> : <i>abhijñā-</i> > <i>avhiñā-</i> , <i>lābha-</i> > <i>lavha-</i> , <i>labhanti</i> > <i>lavheti</i>

		<i>bh > h: aśubha- > aśuha-, ubhaya- > uhae-, śubha- > śuha-</i>
RS-5	<i>bh</i>	<i>bh > h / v: abhi- > abhi- / avi-</i> <i>bh > Ø: aśubha > aśua</i>
RS-12	<i>bh</i>	<i>bh > v: lābhāya > (*la)va; bh > y: abhi- > ayi-</i> <i>bh > bh at compound boundaries: -bhājanataḥ > -bhayaṇaḍa, -bhāvād > -bhave, -bhūmyām > -bhumi</i>
RS-14	<i>bh</i>	<i>bh > v lābhinaḥ > lavina, anabhi- > ana[vi]-</i> <i>bh > h: śobhanam > śohana, -bhiḥ > hi</i>
RS-19	<i>bh</i>	<i>bh > v: abhi- > avi-</i> <i>bh > h: labheya > lahece</i> <i>bh > Ø: BHS abhūṣi > aūṣi, labheya > laeja</i>
RS-20	<i>bh</i>	<i>bh > h: -bhavanti > -bhoti</i> <i>bh > h: -bhavanti > -hoti</i> <i>bh > v: abhi- > avi-, -bhavanti > -veti</i>

The most commonly recorded reflexes of *-bh-* are retention of the historical form, de-occlusion to the aspirate *h*, lenition to *v* (or *vh*), or deletion, which appears to have also been used to represent the weakened forms. Pronunciation of the *h* reflex is assumed to be close to the voiced glottal fricative [h]²¹⁹. The retention of *bh* in the *abhi-* prefix appears to be a common scribal habit (Salomon 2008: 115-116), while *bh* is also retained following prefixes to many forms of the verb *√bhū* (Schlosser 2016: 76), or at compound boundaries (Silverlock 2015: 239). The development of *bh > h* is common in the instrumental plural declension *-bhiḥ > -hi*. The writing of *bh > vh* by BCS5 suggests the retention of aspiration during lenition of the stop component, and staged developments such as *bh > vh > h* (Baums 2009: 145-146), or *bh > [β] > v/vh* (Schlosser 2016: 76), have been suggested.

8.2.1.2.6 Overview – intervocalic stops

A common pattern which emerged through the above sections is that each of the single intervocalic stops is represented by several graphemes in the Kharoṣṭhī record, including historic, transitional and advanced forms. It is likely that these graphemic variants reflect actual variations in pronunciation of the language, whether as retained historicisms, allophones, innovative forms or dialectic variants for example (§ 7.2). This raises questions about the extent to which the Gāndhārī stops had lost or retained

²¹⁹ Baums 2009: 145-146; Silverlock 2015: 238

phonemic contrast in intervocalic positions, and how these changes might be recorded in the language. What is the best representation for each of these phonemes in the lemmas of words - which graphemes provide the most appropriate representation of the underlying phonemes suggested by the orthographic variants? The widespread use of variable forms suggests that the Gāndhārī scribes also attempted to address this problem, and apart from retaining conservative forms in many instances, typically responded in one of two ways in the recording of weakened forms. One approach was to mark the original graphemes with diacritics which indicated lenition to some extent, while the other was to use the graphemes of existing phonemes having similar sounds to the weakened pronunciation.

Common changes to the unaspirated stops are set out in the consonant chart in Table 8.2.1.2.6a, and for the aspirated stops in Table 8.2.1.2.6b, highlighting the different patterns of variant forms. These charts do not include all of the recorded variants, but are intended to highlight the main patterns of change. The most commonly recorded reflexes are highlighted in bold, while those in round brackets occur less frequently. In these charts the OIA stops (T) are listed across the top row, while recorded Gāndhārī stop (T), sibilant (S) and approximant (R) reflexes linked to these stops are shown in the lower rows, with the S and R forms indicating weakened sounds. Reflexes recorded by the scribes using other available graphemes are linked by darker shading, while changes recorded by diacritic marked reflexes are linked by lighter shading.

Table 8.2.1.2.6a: Common changes to single intervocalic unaspirated stops in Gāndhārī

		Labial		Dental		Retroflex		Palatal		Velar	
OIA	T	<i>p</i>	<i>b</i>	<i>t</i>	<i>d</i>	<i>ʈ</i>	<i>ɖ</i>	<i>c</i>	<i>j</i>	<i>k</i>	<i>g</i>
Variant G reflexes	T		(<i>b</i>)		<i>d</i>		<i>ɖ</i>	(<i>c</i>)	(<i>j</i>)		<i>g</i>
	S			(<i>s</i>)	<i>ḍ</i> [ḍ]						<i>ḡ</i> [ɣ]
	R		<i>ɸ</i>						<i>y</i>		(<i>y</i>)
Darker shading: where reflexes include other available graphemes Lighter shading: where reflexes include diacritic marked graphemes Commonly recorded reflexes are highlighted in bold Reflexes in round () brackets occur less frequently											

Most of the changes reflect lenition to weaker consonants in the same place of articulation, as in the case of *bh* > *ɸ* for example, but also involving the merger of

unvoiced with voiced consonants in the majority of cases, as in $\{p, b\} > v$ and $\{c, j\} > y$ for example, while $\{kh, gh\} > h$ also involves a one-step change of place. Assumed pronunciations for the diacritic marked reflexes are indicated in IPA format. The two patterns of development are analysed separately in the following sections.

Table 8.2.1.2.6b: Common changes to single intervocalic aspirated stops in Gāndhārī

		Labial		Dental		Retroflex		Palatal		Velar		Glottal	
OIA	T	(ph)	<i>bh</i>	<i>th</i>	<i>dh</i>	<i>ṭh</i>	<i>ḍh</i>	(ch)		<i>kh</i>	<i>gh</i>		
Variant G reflexes	T		(bh)				<i>ḍh</i>			(kh)	(g)		
	S			<i>s</i>	<i>ś</i> [ṣ]						(g)		<i>h</i> [ɦ]
	R		<i>v, h</i>										
Darker shading: where reflexes include other available graphemes Lighter shading: where reflexes include diacritic marked graphemes Commonly recorded reflexes are highlighted in bold Reflexes in round () brackets occur less frequently													

8.2.1.2.6.1 Reflexes recorded using diacritic markers

Reflexes of the OIA dental and velar unaspirated stops, *t*, *d* and *k*, *g*, are often represented in the Gāndhārī record by the diacritic marked reflexes *ḍ* and *g* respectively, especially by the RS scribe. These modified graphemes are not used consistently, however, and the dental and velar stops are also recorded in their unmodified forms in many cases. The existing grapheme *y* is also used to represent reflexes of the velar stops in some instances, and the phonemes are sometimes omitted by the RS scribe, as in $t > \emptyset$ for example. This overall pattern suggests that although the scribes were aware of the lenition of these sounds, they may have been uncertain how best to represent these consonants in view of the range of variations in pronunciation to which they were exposed. In this situation it is reasonable to conclude that these stops retained a distinct phonemic status in intervocalic positions, and the modified forms represent attempts to capture or mark allophonic pronunciation in intervocalic environments. And while it is possible that the unaspirated velar stops may have reverted to the palatal *y* in pronunciation, this grapheme is used less often than it is for the palatal stops (§§ 8.2.1.2.1, 8.2.1.2.2), and *g*, *ḡ*, *k* and *ḱ* are the most common forms used for the velars. This suggests that the velar place was

retained, while *y* may have simply been adopted on occasion as a convenient grapheme whose pronunciation was similar to an allophone such as the fricative [ɣ]. While the use of orthographic *y* along with the historic forms for both *k* and *g* commenced in the Aśokan inscriptions, the introduction of diacritic marked reflexes to capture variable pronunciations only occurred in the later stage first and second century CE manuscripts, suggesting the perception of more transitional pattern of sound changes than could be captured with the available phonemic forms, supporting the above interpretation.

As discussed in § 7.3, two approaches might be followed in selecting a best representation for these phonemes in the lemma forms of words. One is to retain the historic forms, while the other is to adopt some form of transitional representation. In the first approach, in the absence of clear evidence that this was lost, the orthographic status of the original phonemic grapheme is retained, and all of the other reflexes are recognised as attempts to record particular allophones. In the second case, some form of transitional representation might be introduced, such as *ḍ* and *ḡ*, similar to the RS scribe diacritics. This entails the introduction of new forms in the phonemic inventory, however, which are in reality only distinct from the word initial stops at an allophonic level rather than as phonemes. While such markings are potentially useful in avoiding graphic confusion created by the use of single stop phonemes to represent geminates and other reduced clusters, this is not a problem in modern transliterations where these are normally recorded as actual geminates (§ 8.3). Since voicing of the unvoiced stops *t* and *k* is well established, the following representation of the stop pairs in the development from Skt. > G is suggested: (*t*, *d*) > *d* / V_V, (*k*, *g*) > *g* / V_V.

Single retroflex unaspirated and aspirated stops are not very common in OIA, and on the evidence of a small data sample these phonemes appear to have been retained. While it is also possible there was insufficient impetus to address any weakened forms in the Kharoṣṭhī script, the following representations are suggested, which incorporate an assumed voicing of the unvoiced stops: (*t*, *ḍ*) > *ḍ*, and (*th*, *ḍh*) > *ḍh*.

8.2.1.2.6.2 Reflexes recorded using the graphemes of other phonemes

Reflexes of the OIA labial and palatal unaspirated stops, *p*, *b* and *c*, *j*, are mostly represented in the Gāndhārī record by the existing phonemes *v* and *y* respectively, while the labial aspirated voiced stop *bh* is commonly represented as *v*, the dental aspirated

stops, *th*, *dh*, are often represented by *s* or the diacritic marked *ṣ*, and the velar aspirated stops *kh*, *gh* are commonly represented by *h*. Alternative reflexes including unmodified forms are sometimes used however, such as *bh* > (*bh*, *v*) for example. The overall pattern suggests that these intervocalic stops were weakened to the extent that they sounded similar to the approximants. It is unclear, however, whether this usage represents a full merger between the sounds of the phonemes, or was possibly adopted as a convenient marker of the weakened sounds, which sounded similar to, but not necessarily the same as the approximants. This interpretation is supported by the use of alternate forms such as *kh*, *gh* > *g* and *th*, *dh* > *ṣ* by the RS scribe, and *bh* > (*bh*, *v*) by all scribes, suggesting that some uncertainty remained about how to best represent these phonemes. Hence these reflexes may not represent a full phonemic merger similar to those that occurred at later stages in other MIA languages (§ 9.1.1), but are rather an indication of a partially developed but not fully advanced sound change in progress.

These changes also raise the question of whether the two sets of sounds had merged sufficiently for homonym pairs to emerge, and if so, how were these dealt with in the language. The development of homophonous or homonymous forms, which have the same sound but different or the same spellings respectively, is common in language change, and languages adjust for such developments in different ways, such as through lexical replacement or loss for example (Campbell 2016: 330-332). The extent to which homonym pairs may have developed in Gāndhārī is difficult to assess, even with a full lexical analysis, let alone how these may have been absorbed or adjusted for within the language at the time. In a synthetic language such as Sanskrit many quite complex words are formed which may include a prefix or preverb, a stem built up from a root plus one or more affixes, along with a declensional or conjugational ending. And while phonemic mergers may lead to loss contrasts in one or several of these morphemes, the extent to which these may affect the overall word meaning is difficult to assess. This can be seen in the following examples:

- Skt. *abhi-muñc-a-ti* (*abhi* + $\sqrt{\text{muc}}$ [“release” MW. s.v.] present tense, 3rd person singular) > G *avi-muc-a-di* (Glass 2007: 117). The Sanskrit form is made up from the Sanskrit preverb *abhi-*, the modified verbal root *muñc*, a present stem affix *-a-*, plus the conjugational ending *-ti*. The verb root *muñc* > *mu(ñ)c* is effectively unchanged in Gāndhārī even though the nasal is not recorded (§ 8.3.2.1), and any homonymous morphemes appear to be associated with the preverb affix. For example, the Sanskrit

preverb *api-* (“in, near” MW s.v.) might also be rendered as *avi-* in Gāndhārī, and although not often used in this context, lends a slightly different meaning to the verb than *abhi-* (“to, towards” MW s.v.). It is possible however that such differences could either merge or be understood contextually.

- Skt *abhi-kāṅkṣ-at-ā* (*abhi* + $\sqrt{kāṅkṣ}$ [“desire” MW s.v.] present participle, masculine instrumental singular) > G *abhi-ghakṣ-ad-a* (Salomon 2008: 115, 424). The Sanskrit form is made up from the preverb *abhi-*, the verbal root *-kāṅkṣ-*, a suffixal *-at-* to form the participle, and a final long *-ā* declensional ending. The development of the verb root *kāṅkṣ* > *gaks* (*ghaks*) does not appear to create a homonymous form (MW s.v.), while the preverb *abhi-* has been retained in this case, following the common scribal habit in Gāndhārī (§ 8.2.1.2.5). While this is assumed to reflect the retention of historic spellings in conservative environments, this could also reflect a desire to retain a form which contrasts with the reflexes of other preverbs such as *api-*. This suggests some form of sound contrast was possibly retained, or the scribes may have preferred to retain a lexical contrast in homophonous forms for greater clarity in the written texts.

It seems unlikely that such complex verbal forms would arise in Gāndhārī from different morphological derivations, and overall interpretation and meaning do not appear to be affected significantly by homonymous morphemic components within the word. It is also possible that finer distinctions of meaning such as between *abhi-* or *api-* > *avi-* were lost or accepted during the diachronic development of such sound changes. Similar patterns occur in the reflexes of simpler words, as seen in the following examples:

- Skt. *sukha-* (adj. “comfortable, pleasant” MW s.v.) > G *suha-* (Salomon 2008: 109), which does not appear to create a homonymous form.
- Skt. *rājan-* (m. “a king, sovereign” MW s.v.) > G *raya-* (Salomon 2008: 109) could become homonymous with Skt. *raya-* (m. “current, stream” MW s.v.) > G *raya-*.
- The uncommon development of Skt. *nakha-* (m. “nail, claw” MW s.v.) > G *ṇaga-* (Glass 2007: 115) may become homonymous with developments of Skt. *naga-* (m. “mountain, tree” MW s.v.) and *nāga-* (m. “serpent” MW s.v.).

The general breakdown of conjugational and declensional endings in the MIA languages (§ 2.2.3) also occurred in Gāndhārī, leading to the loss or merger of many case endings (e.g., Baums 2009: 212-213). This points towards a general diachronic adjustment

for homonymous forms within the language, and it is likely that the word medial changes to intervocalic consonants were also absorbed into the language in similar ways.

Unlike the diacritic marked reflexes, it is difficult to establish a clear approach to the best form of representation for these reflexes, which are recorded using the graphemes of other phonemes, to be used in the lemma forms of words. On the one hand, it is evident that the sounds of these stops were in the process of merging with the homorganic sibilants and approximants during the early centuries CE. On the other hand, it remains uncertain whether these sounds had merged fully, or if they had, whether this was accepted without qualification by the scribes. In this context the phonemes might be represented in one of several ways. One approach is to retain the original forms and record the variants, but this falls short of reflecting the extent of change that has occurred. Another is to simply record the new merged phonemic forms and accept that the language had adjusted to any new homonymous forms, which is similar to the full phonemic merger suggested by Baums (2009: 143, 147). However, as observed previously (e.g., §§ 1.4.4, 7.2), the range of variant reflexes most likely reflect actual variations that existed or remained in the pronunciation of the language, whether in the main or other Gāndhārī dialects, and it seems appropriate to represent the uncertain and apparently incomplete status of the mergers in some way. This leads to an approach in which some form of diacritical marking is introduced to reflect the uncertainty of this situation, which might be included in two ways:

- Marking the original OIA consonant: $(p, b) \equiv \underline{b} [b]$, $bh \equiv \underline{bh} [b^h]$, $(th, dh) \equiv \underline{dh} [d^h]$, $(c, j) \equiv \underline{j} [j]$, $(kh, gh) \equiv \underline{gh} [g^h]$. This system represents a more conservative transitional approach in that it records both the original OIA phoneme while indicating that it has been weakened by the addition of a diacritic marker. The sound changes are recorded in a similar way without indicating a specific sound reflex.
- Marking the apparently merged OIA consonant: $(p, b) \equiv \bar{v} [v]$, $bh \equiv \bar{h} [h]$, $(th, dh) \equiv \tilde{s}$, $[z]$, $(c, j) \equiv \bar{j} [j]$, $(kh, gh) \equiv \bar{h} [h]$. This system represents a more advanced approach in that it records the assumed merged OIA phoneme, while also indicating that it has developed from another phoneme by the addition of a diacritic marker, which also suggests that its pronunciation may also vary. The use of \tilde{s} is to distinguish from \bar{s} which is used often in reflexes of SC clusters (§ 8.3.3). The assumed sounds are recorded in a similar way.

While more conservative, the first approach seems to offer a more appropriate representation. In this approach the OIA source phoneme is recognised, while the nature of the progressive sound shift is also indicated. An interesting comparison can be made with the known phones and phonemes of the modern day Torwali language, where many of the major classes of stops appear to have retained phonemic status in intervocalic positions. At the same time the pronunciation is weakened in many cases, similar to that assumed for Gāndhārī, such as the pronunciation of *b* which is weakened to the allophone [v] in intervocalic positions (Lunsford 2001: 11-26). Such situations create difficulties in the phonemic analysis of the language, and the scribes in ancient Gandhāra must have faced similar transcriptional problems. Faced with the dilemma of recording a changing language in a script that was not fully adapted for this purpose, and mostly lacking a clear understanding of the distinction between a phoneme and its expressed allophones, a degree of phonetic and phonemic confusion appears to have ensued among the scribes.

8.2.1.2.6.3 Summary and phonemic representation

The changes discussed in the previous section and represented in Tables 8.2.1.2.6a and 8.2.1.2.6b indicate variable lenition of the intervocalic stops depending upon their place of articulation. The weakening appears to have progressed close to a full sound merger with the approximants *ɣ* and *ʏ* in the case of the unaspirated labial and palatal stops, and with the voiced glottal fricative *h* in case of the aspirated velar and labial stops. Lenition of the unaspirated dental and velar stops along with the aspirated dental stop appears to have progressed to a lesser extent to a fricative pronunciation. In the earlier Aśokan inscriptions, reflexes such as *j* > {*j*, *ɣ*, *c*}, *dh* > *dh*, *p* > {*p*, *ɣ*} and *bh* > {*bh*, *h*} are also recorded. While these indicate that sound changes were occurring, the historic forms appear to be retained more often, suggesting that the sound changes were possibly less advanced at that time. The inclusion of advanced forms may also reflect copying from the Māgadhī Prakrit of Aśoka in which the inscriptions were originally composed (§ 2.2.5).

The changes in Gāndhārī contrast with developments recorded in the later MIA languages, as represented in the Mahārāṣṭrī and Śaurasenī Prakrits for example, in which, with the exception of the retroflex stops for both MIA and Gāndhārī, lenition and merger of all of the unaspirated stops with the palatal approximants *ɣ* or *ʏ* occurred, with resultant elision in some cases, along with merger of all of the aspirated stops with the glottal

fricative h^{220} . At the same time all of these stops are typically preserved in Pali (Oberlies 2019: 114-115), and it has been suggested that the changes which occurred in Gāndhārī represent a so-called “middle stage of development” of the intervocalic stops in the MIA languages (Salomon 1999: 124-125). While languages such as Pali and Mahārāṣṭrī present a reasonably consistent record of either rarely occurring or complete change, the range of variant reflexes in Gāndhārī most likely record actual variations that existed or remained in the pronunciation of the language, whether in the main or other Gāndhārī dialects. This variable recording of the changes in progress in the Gāndhārī record, including transitional sounds, and the various orthographic techniques used to record these, provide a valuable record of language change in progress. Based on the observed reflexes it seems clear that the voiced - voiceless distinction was lost in favour of voicing, and that there was a general tendency, of changing the voiced stop outcomes to fricatives or approximants, seemingly representing a change in progress.

Summary: A summary of the suggested phonemic representations of Gāndhārī reflexes of OIA intervocalic stops in the lemma forms of words is presented in Table 8.2.1.2.6.3. This includes listings of both conservative and advanced approaches to the representation of phonemes as discussed in the previous section. The conservative representations are based on the original OIA graphemes, but with the addition of diacritic marks in those cases which appear most advanced towards approximants in the Gāndhārī reflexes. Advanced form options are also listed for these cases, reflecting the potential merged phonemes. These suggested representations can be varied to suit different interpretations or additional data which may become available, while the use of particular reflexes may be word dependent. The variant reflexes recorded in the Kharoṣṭhī source materials for each phoneme are also listed and classified according to their assessed status.

8.2.1.3 Word final

While it is implicitly understood in the source publications that OIA word final consonants are lost in Gāndhārī, this is not explicitly stated by scholars in most cases, with some exceptions like Hultzsch (1922: lxxxvii) for example, presumably because this is the general case with all of the Prakrits.

²²⁰ E.g., Misra 19.67: 169-170; Masica 1993: 180-182; Salomon 1999: 124-125

Table 8.2.1.2.6.3 Suggested representation of Gāndhārī reflexes of OIA intervocalic stops in the lemma forms of words

OIA Stop	Suggested G phoneme		Assumed sound (option)	Alternate representations of Gāndhārī phonemes recorded in the source materials (≡)
	Conservative representation	Advanced representation		These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
<div><div>/-k-/</div><div>/-g-/</div></div>	<div><div>/-g-/</div><div>/-g-/</div></div>	<div><div></div><div>/-Ø-/</div></div>	<div><div>[q]</div><div>([ʏ] [j])</div><div>([j])</div></div>	<div><div>/g/ ≡ {/g^h, /gh^{alt}, /g^{tr}, /Ø^{adv}, (/k^h, /k^{tr}, /y^{adv})}</div><div>/g/ ≡ {/g^h, /gh^{alt}, /g^{tr}, /Ø^{adv}, /y^{adv}, (/k^{tr})}</div><div>Elision occurs in the environment of at least one high front vowel (<i>i</i> / <i>e</i>).</div></div>
<div><div>/-kh-/</div><div>/-gh-/</div></div>	<div><div>/-gh-/</div><div>/-gh-/</div></div>	<div><div>/-h̄-/</div><div>/-h̄-/</div></div>	<div><div>[q^h]</div><div>([ḥ])</div></div>	<div><div>/gh/ ≡ {/h^{adv}, (/kh^h, /k^{tr}, /Ø^{adv}, /g^h)}</div><div>/gh/ ≡ {/h^{adv}, /g^h, (/g^{tr}, /Ø^{adv}, /kh^h)}</div></div>
<div><div>/-c-/</div><div>/-j-/</div></div>	<div><div>/-j-/</div><div>/-j-/</div></div>	<div><div>/-ȳ-/</div><div>/-ȳ-/</div><div>/-Ø-/</div></div>	<div><div>[ɟ]</div><div>([j])</div><div>([j])</div></div>	<div><div>/j/ ≡ {/y^{adv}, /Ø^{adv}, (/c^h, /j^h)}</div><div>/j/ ≡ {/y^{adv}, (/Ø^{adv}, /j^h, /c^h)}</div><div>Elision occurs in the environment of at least one high front vowel (<i>i</i> / <i>e</i>).</div></div>
<div><div>/-ṭ-/</div><div>/-ḍ-/</div></div>	<div><div>/-ḍ-/</div><div>/-ḍ-/</div></div>	<div><div>/-ḍ-/</div><div>/-ḍ-/</div></div>	<div><div>[ɽ]</div></div>	<div><div>/ḍ/ ≡ {/ḍ^{adv}, (/ṭ^h)}</div><div>/ḍ/ ≡ /ḍ^h (limited examples)</div></div>
<div><div>/-ṭh-/</div><div>/-ḍh-/</div></div>	<div><div>/-ḍh-/</div><div>/-ḍh-/</div></div>	<div><div>/-ḍh-/</div><div>/-ḍh-/</div></div>	<div><div>[ɽ^h]?</div></div>	<div><div>/ḍh/ ≡ /ḍh^h (single example)</div><div>/ḍh/ ≡ /ḍh^h (limited examples)</div></div>
<div><div>/-t-/</div><div>/-d-/</div></div>	<div><div>/-d-/</div><div>/-d-/</div></div>		<div><div>[ɖ]</div><div>([ḍ̥])</div></div>	<div><div>/d/ ≡ {/d^h, /d^{tr}, (/ḍ^{alt}, /Ø^{alt}, /s^{alt}, /ṣ^{alt})}</div><div>/d/ ≡ {/d^h, /d^{tr}, (/Ø^{alt}, /t^{alt}, /ḍ^{alt})}</div></div>
<div><div>/-th-/</div><div>/-dh-/</div></div>	<div><div>/-dh-/</div><div>/-dh-/</div></div>	<div><div>/-ṣ̣-/</div><div>/-ṣ̣-/</div></div>	<div><div>[ɖ^h]</div><div>([ṣ̣])</div></div>	<div><div>/dh/ ≡ {/ṣ^{adv}, /s^{alt}, (/th^h, /dh^{alt}, /ḍ^{alt})}</div><div>/dh/ ≡ {/ṣ^{adv}, /s^{alt}, (/dh^h)}</div></div>
<div><div>/-p-/</div><div>/-b-/</div></div>	<div><div>/-b-/</div><div>/-b-/</div></div>	<div><div>/-v̄-/</div><div>/-v̄-/</div></div>	<div><div>[b̥]</div><div>([v̄])</div></div>	<div><div>/b/ ≡ {/v^{adv}, /Ø^{alt}, /p^h, (/b^{alt})}</div><div>/b/ ≡ {/v^{adv}, /Ø^{alt}, (/b^h)}</div></div>
<div><div>/-bh-/</div></div>	<div><div>/-bh-/</div></div>	<div><div>/-h̄-/</div></div>	<div><div>[b^h]</div><div>([ḥ])</div></div>	<div><div>/bh/ ≡ {/h^{adv}, /bh^h, /v^{alt}, (/Ø^{alt}, /vh^{alt}, /y^{alt})}</div></div>

Word final stops are relatively infrequent in OIA, and the loss of stops and other consonant endings is mostly seen in the reflexes of OIA consonant final case endings²²¹. These occur in the nominal *-a* stem masculine and neuter ablative singular ending, *-āt*, and the nominative singular of consonant ending stems such as *vāk* (MW. s.v. $\sqrt{v}āc$ f “speech”) and *marut* ($\sqrt{m}arut$ m “wind”) for example, while word final *-t* also occurs in the imperfect 3rd person singular verbal suffix, as well as some pronouns such as *tad*. In the Sanskrit sandhi rules, word final stops typically form concatenated clusters with following word initial consonants, where they may be altered by the regressive assimilation of some features (§ 4.2). In ‘sentence’ final positions where they remain unjoined, the consonants are marked with a virāma diacritic, which indicates the absence of a final vowel. The following examples of Gāndhārī ablative singular endings written by various scribes show a similar development to those seen in Pali:

- G *abhaviḍatva cadoṇa* for assumed P *abhāvittā cattunnam* or Skt. *abhāvitatvāt caturṇām* (Glass 2007: 128, 137, 241, 243. [RS scribe]). In Sanskrit sandhi *-āt cat-* would be written as *-āccat-*.
- G *padātala tvayapayata* for assumed P *pāda-talā taca-* or Skt. *pāda-talāt tvaca-* (Glass 2007: 128, 135, 244, 245. [RS scribe]). In Sanskrit sandhi *-āt tva-* would be written as *-āttva-*.
- G *sakṣeva jīamaṇa* for assumed P *saṅkhepā jīyāmāṇam* or Skt. *saṅkṣepāt jīryāmāṇām* (Baums 2009: 208, 248, 641, 678. [BL scribe 4]). In Sanskrit sandhi *āt jīrya-* would be written as *-ājīrya-*.
- G *dukhaspa di* for assumed P *dukkhasmā ti* or Skt. *duḥkhāt iti* (Glass 2007: 128, 137, 245. [RS scribe]). In Sanskrit sandhi *-āt iti* would be written as *-āditi-*.

There are however some examples where historical final consonants in pronouns preceding vowels are retained in set expressions such as P *yāvad eva* > G *yavad eva* and *etad avaca* > *eghad oya* (Allon 2001: 101).

A number of other forms involving different vowel finals in reflexes of Sanskrit ablative *-āt* ending have been observed in some manuscripts. This variable orthography does not appear to follow any consistent linguistic pattern (Silverlock 2015: 300-301), and may reflect a similar variability to the development of word final vowels in the direct case

²²¹ E.g., Allon 2001: 106-118; Glass 2007: 126-134; Baums 2009: 201-239; Schlosser 2016: 91-101.

endings (Baums 2009: 201-213). A diacritic similar to the Sanskrit virāma, which indicates a final non-syllabic consonant, has only been observed in the Niya Documents and is not apparent in the Gandhāran Kharoṣṭhī record (Glass 2000: 138). This may not have been developed or used for what was possibly only an occasional need.

Summary: OIA word final stops are generally deleted in Gāndhārī, as in other Prakrits, and the development from Skt. > G is typically represented as /-VT#/ > /-VØ#/.

8.2.2 Nasals

8.2.2.1 Word initial

Only the labial and dental nasals *m* and *n* occur word initially in OIA (§ 4.4), and appear to remain stable in these positions in Gāndhārī, although *n* is frequently written as *ṇ*²²², depending upon the scribe. One exception noted is the dissimilation of *m* > *v* seen in Skt. *mīmāṃsā* > G *vimasa* recorded in the Nid-G, (Baums 2009: 138-139). Similar dissimilated forms are found in other MIA languages, suggesting a common innovation. The palatal nasal *ñ*- is recorded word initially in Gāndhārī, but this only arises as a reflex of the OIA word initial *jñ*- cluster (§ 8.3.1.2).

8.2.2.2 Word medial

The nasals *m*, *n*, and *ṇ* occur independently in word medial positions in OIA, while *ñ* and *ṇ* occur almost entirely in homorganic nasal + stop clusters, where they are essentially non-contrastive allophones (§ 4.4). Word medial *m*, *n* and *ṇ* are recorded as single intervocalic consonants in the Kharoṣṭhī script, and although medial *ñ* is also written as a single consonant, this only arises as a reflex of OIA *-jñ*- and *-{n, ṇ}y*- clusters and is considered to represent the geminate nasal *-(ñ)ñ*- (§§ 8.3.1.2 and 8.3.2.4.3). The velar nasal *ŋ* is not recorded in the Kharoṣṭhī script.

8.2.2.2.1 The labial nasal *m*

The labial nasal *m* typically remains stable in intervocalic positions²²³. Two examples of the apparent deletion of *m* are recorded in the EĀ-G in Skt. *arciṣmanti* (P

²²² Konow 1929: xcvi; Mehendale 1948: 12, 19; Salomon 2000: 81; Allon 2001: 78; Glass 2007: 114, 117; Salomon 2008: 106; Baums 2009: 137-139; Schlosser 2016: 71.

²²³ Mehendale 1948: 13; Konow 1929: civ; Allon 2001: 79, 85; Glass 2007: 114, 117; Baums 2009: 137.

accimantāni) > G *aceata* and Skt. *caṣṣuṣmantah* (P *cakkhumanto*) > *caṣṣu[a]tu* (Allon 2001: 85). These both involve the addition of the *-mant* suffix to *-is/-us* stems in Sanskrit leading to the formation of *-ṣm-* clusters, in comparison to *-i/-u* stems and intervocalic *-m-* recorded in Pali. An interchangeability between *-m-* and *-v-* has been observed in the Dhp-G^K for example²²⁴, although these could be interpreted as borrowed forms with the glide *-v-* left unwritten. This development differs from the *-ṣm-* > *-Sp-* reflexes observed for other Sanskrit clusters (§ 8.3.3.2), supporting the interpretation of borrowed forms.

8.2.2.2.2 The dental and retroflex nasals *n* and *ṇ*

While similarities exist between the akṣaras for *n* and *ṇ* as vertical strokes with a slanted mark to the right at the top (Table 5.1), which can lead to graphic confusion in a loose scribal hand, they retain sufficient contrast in the formation of the top marking as a curved or sharp hook to be identified as distinct graphemes in most documents. The Gāndhārī reflexes of *n* and *ṇ*, recorded in the various source materials are discussed below.

Aś-H, Aś-M (3 BCE): Sanskrit *ṇ* is generally recorded as such, although some exceptions occur, such as Skt. *kaṣamaṇīya* > G *kaṣamanaye*, and it is not recorded in some terminations following *r* or *ṣ* where it otherwise occurs in Sanskrit, such as G *putrena*, *vagreṇa* and *sakasani*. This suggests that retroflexion of *n* according to the NATI rule was no longer being strictly observed in Gāndhārī. Sanskrit *n* also appears to be preserved when not subject to NATI assimilation²²⁵.

Ins-K, Ins-M (1 BCE – 3 CE): Konow (1929: cii–civ) presents a number of examples which show varied preservation and change of the two nasals, and even interchange of the two symbols, for example for Skt. *śākyamuni* > G *śakamuṇi* and *jana-* > *jaṇa-*. It is suggested that intervocalic *n* and *ṇ* had the same sound, at least over most of the territory, which was probably cerebral, while the traditional orthography restrained the use of a consistent script. Mehendale (1948: 299) observes that cerebral *ṇ* was strongly preserved in the first century BCE, and mostly in the first century CE, but the reversion to dental *n* appears to have become more common in later centuries, as for example in Skt. *kaṣuṇa* > G *kaṣuna*, *nirvāṇa-* > *nirvana-* (2 CE), and *dakṣiṇa-* > *dakṣina-* and *śramaṇa-* > *ṣamana-* (3 CE). While dental *n* is preserved in many cases, cerebralisation is reported in

²²⁴ Konow 1929: civ; Allon 2001: 85.

²²⁵ Hultsch 1922: lxxxv; Mehendale 1948: 19-20, 212, 222. The NATI assimilation rules are described in Appendix A1.

examples given from the first century BCE, as in Skt. *bahujana-* > G *bahujāṇa-*, through to the fourth century CE, as in *tanaya* > *taṇaya-* for example (Mehendale 1948: 304-305).

Gandhāran Buddhist Manuscripts (1 CE – 2 CE): The orthographic treatment of *n* and *ṇ* varies between the scribes. The following scribes typically used orthographic *ṇ* exclusively for both *n* and *ṇ*.

- BL Scribe 1: Represented in BLS1-1 (Salomon 2008: 96), BLS1-2 (Allon 2001: 69), and BLS1-2 (Lenz 2003: 41). Examples include Skt *anavatapte* > G *aṇodate* (Salomon 2008: 424), *kāyena* > *kayeṇa* (Allon 2001: 328), and *nīvaraṇe* > *ṇīvaraṇa* (Lenz 2003: 259).
- BL Scribe 4: Represented in BLS4-1 (Baums 2009: 134-135). Examples include Skt. *anupādānaḥ* > G *aṇ[u]adaṇo* and *nidānānām* > *ṇidaṇaṇa* (Baums 2009: 622, 650).
- BL Scribe 9: Represented in BLS9-1 (Salomon 2000: 75). Examples include Skt. *āsevamāna-* > *asevamaṇa* (Salomon 2000: 228).
- BCS5-1 Scribe 5: Represented in BCS5-1 (Schlosser 2016: 47). Examples include Skt. *darśana* > G *darśaṇa* and *naiva* > *ṇeva* (Schlosser 2016: 302,
- RS Scribe: Represented in RS-5 (Glass 2007: 107, 115-116) amongst others. Examples include Skt. *nakhāḥ* > G *ṇaga* (Glass 2007: 116).

In contrast to these scribes, BL Scribe 2 typically used only the *akṣara* for dental *n* and not retroflex *ṇ*:

- BL scribe 2: Represented in BLS2-1 (Lenz 2003: 119-120) and BLS2-2 (Lenz 2010: 18, Table 3). Examples include Skt. *ātīrṇaḥ* > G *adīrno* (Lenz 2003: 129) and *grhīṇati* > *ghin[adi]* (Lenz 2010: 26).

While the orthographic distinction between the writing of dental *n* and retroflex *ṇ* appears to have been largely maintained in the third century BCE Aśokan inscriptions, a number of contrary examples suggest that the language was at that stage beginning to vary from Sanskrit concerning these nasals. In later usage from around the first century CE the two sounds were typically recorded using either one or the other of the characters, which appeared to depend upon scribal preference. This suggests an appreciation that there was very little if any lexical distinction between words spelt one way or the other, which is consistent with the already limited phonemic status of *ṇ* in Sanskrit (§ 4.3). While the use

of both symbols has been interpreted as an indication that a levelling or merging of the sounds of two phonemes had occurred²²⁶, an alternate interpretation is that a distinct retroflex pronunciation was still evident in particular environments, but its allophonic nature and lack of phonemic contrast in minimal pairs was also appreciated. The patterns of use also suggest that the scribes were attempting to determine a most appropriate underlying form to use in representing the *n* phoneme (Davenport & Hannahs 2010: 123), and were not necessarily interested in or informed by the Sanskrit sandhi system. A number of factors point towards the selection of dental *n* rather than retroflex *ṇ* as the most appropriate underlying form: it occurs more frequently in word medial positions and is used consistently word initially in Sanskrit, while also having a phonetically simpler articulation. At the same time the common occurrence of conditioned *ṇ* suggests it was afforded an equal importance by many scribes. It is interesting that orthographic *n* was adopted by BL scribe 2 in the recording of avadāna and pūrvayoga texts, which may bear a closer relationship to the spoken language than other more formal texts (§ 3.3). This is not seen in the scholastic texts recorded by BC scribe 5, however, where orthographic *ṇ* was adopted.

8.2.2.3 Word final

While the nasals *m*, *n*, *ṁ* and *ṇ* occur word finally in Sanskrit, these are reduced to *m*, *n* and *ṇ* and assimilated in different ways through the external sandhi rules at word junctions. Similar to the omission of word final stops (§ 8.2.1.3), these nasals were omitted in most instances in the Kharoṣṭhī record. This is seen for example in the masculine accusative singular endings for Sanskrit *-a* stems, where *-am* is reduced to *-a* and *-o* in the majority of cases, such as Skt. *ayam* > G *aya* (Hultzsch 1922: lxxxvi), *udāram* > *uraḍa* and *nipakam* > *ṇivago*, while *putram* > *pat[r]am* is also recorded (Salomon 2000: 95, 228, 230, 231), *dharmam* > *dhama*, *śabdān* (masculine accusative plural) > *chado* (Allon 2001: 109, 331, 329), and *grāmam* > *gama* (Glass 2007: 127, 243), to record only a few of the extensive lists of examples in the source publications. Similarly in the genitive plural endings where *-ām* is reduced to *-a*, such as *gurūṇām* > *guruna* (Hultzsch 1922: lxxxvi), *āsravānām* > *asavaṇa* (Glass 2007: 129, 241), and **anumadhyānām* > *aṇumajāṇa* (Baums 2009: 211, 623). Similar to the Devanāgarī script, *anusvāra* is recorded as a diacritic mark on the preceding syllable in Kharoṣṭhī, where it

²²⁶ Salomon 2000: 75; Glass 2007: 115-116; Baums 2009: 135

indicates the post-vocalic nasalisation within that syllable (§ 5.3.2.3). In this context it would have been quite straightforward for the scribes to record any word final *anusvāra* within the prior syllabic *akṣara*, providing support to the observation that Sanskrit word final nasals were omitted in Gāndhārī.

It has been suggested that the omission of word final *anusvāra* in Gāndhārī may have been due to a weakened articulation of the nasal sound as much as scribal laxity (Hultzsich 1922: lxxxvi). Based on extensive analysis of the *-a* stem developments, Baums (2007: 213-213) suggests that although not written, a nasal segment was retained phonetically following short vowels, while it was omitted following long vowels, at least in the early stages of the recorded language, and this nasalisation appeared to have weakened as the language developed further. Nasalisation of the preceding vowel is recorded in some of the non-Kharoṣṭhī major pillar and rock edicts, as for example in Skt. *dānam* > *danam* (MIA) (Mehendale 1948: 21), and is also recorded in Pali and other later Prakrits²²⁷.

8.2.3 Fricatives

8.2.3.1 Word initial

Sanskrit word initial fricatives generally remain stable in Gāndhārī²²⁸. A small number of atypical developments have been observed:

- The development of BHS *śirṣa-* > G *ṣiṣa-* suggests anticipation of the following *ṣ* (Lenz 2010: 27), or may have developed through metathesis and the influence of the *r*, as in *śir-* > *śri-* > *ṣi*.
- The distant dissimilation of palatal *ś-* > *s-* when this occurs prior to the palatal stop *c*, seen in OIA *śucika* > G *suyia* and *śuci-* > *suyi* for example (Baums 2009: 186-187).
- Processes of remote progressive and regressive palatal assimilation of dental sibilants in both word initial and medial positions are seen in examples such as OIA *śaṃset* > G *śaśea* and *śaśaya* > *śaśaga* (Baums 2009: 186-187), and *śāsana-* > *śaśane* (Salomon 2008: 117), which appear to be regular developments in the language.

²²⁷ Oberlies 2019: 187; Pischel 1957: 240, (§ 349).

²²⁸ E.g., Mehendale 1948: 15,16, 216, 302-303; Allon 2001: 78-79; Glass 2007: 113-114; Baums 2009: 137-139; Silverlock 2015: 222; Schlosser 2016: 71, Marino 2017: 116-117.

While no examples appear to have been listed in the source documents, it is assumed that, similar to the stops (§ 8.2.1.1), when occurring in the second members of compounds word initial fricatives mostly remain stable. Word initial sibilants are also recorded as reflexes of word initial *Sr*- clusters in some cases, although these may also be in borrowings from source dialects. (§ 8.3.3.4.4).

Summary: OIA word initial fricatives generally remain stable in Gāndhārī, and the development from Skt. > G is typically represented as /S_x-/ > /S_x-/, where x is representative of the same place feature. Some exceptions occur which relate to the assimilation or dissimilation of place features, as noted in the above bullet points.

8.2.3.2 Word medial

The Gāndhārī reflexes of OIA single intervocalic fricatives recorded in the Aśokan inscriptions and Gandhāran Buddhist manuscript sources are presented in Table 8.2.3.2a. The alternative forms of the Gāndhārī reflexes of OIA single intervocalic fricatives recorded in the source materials are summarised in the following formulae.

The palatal sibilant *ś*:

Aś Ins: *ś* > {*ś*, (*s*)} / V_V

BL Scribes: *ś* > {*ś*, (*h*), (*Ø*)} / V_V

RS & BCS5 Scribes: *ś* > {*ś*, *ṣ́*, (*s*), (*ṣ́*), (*h*), (*ṣ́*)} / V_V

The retroflex sibilant *ṣ*:

Aś Ins: *ṣ* > {*ṣ*, (*ś*), (*s*)} / V_V

BL Scribes: *ṣ* > *ṣ* / V_V

RS & BCS5 Scribes: *ṣ* > {*ṣ*, *ṣ̣*, (*ṣ́*)} / V_V

The dental sibilant *s*:

Aś Ins: *s* > *s* / V_V

BL Scribes: *s* > {*s*, (*h*), (*Ø*)} / V_V

RS & BCS5 Scribes: *s* > {*ṣ́*, (*s*)} / V_V

The glottal fricative *h*:

Aś Ins: *h* > {*h*, (*Ø*)} / V_V

BL Scribes: *h* > {*h*, (*Ø*), (*s*), (*ṣ́*)} / V_V

RS & BCS5 Scribes: *h* > {*h*, *Ø*, *ṣ́*, *ṣ̣́*, (*ṣ́*), (*v*, *k*, *d*)} / V_V

Table 8.2.3.2a: Gāndhārī reflexes of OIA intervocalic fricatives – all sources

Era	3 BCE	1 CE					
Source:	Aś-M	BLS1-1	BLS1-2	BLS1-3	BLS2-1	BLS2-2	BLS4-1
pages:	15-16, 216	107-117	79-86	40-43	128-130	25-28	137-150
OIA -ś-	ś, (s)	ś, (h), (Ø)	ś	(?)	ś	ś, (Ø)	ś
-ṣ-	ṣ, (ś), (s)	ṣ	(?)	(?)	ṣ	(?)	ṣ
-s-	s, (ś)	s	s	s, (h), (Ø)	s, (h)	s	s
-h-	h, (Ø)	h, (Ø)	h, (s)	h, (Ø)	h	h, (Ø)	h, (ś)
Era	1 CE	2 CE					
Source:	BLS9-1	BCS5-1	RS-5	RS-12	RS-14	RS-19	RS-20
pages:	81-87	72-80	114, 118	229, 244	338, 340	55	125-127
-ś-	ś	ś, (ś)	ś, (s), (ṣ)	ś, (h), (ṣ), (s)	ś	(?)	(?)
-ṣ-	ṣ	ṣ	ṣ, (ṣ)	ṣ, (ś)	ṣ	(?)	ṣ
-s-	s, ṣ	ṣ	ṣ	s, (ṣ)	s	(?)	s, ṣ
-h-	h	ś, ṣ	h, Ø, (ś)	h, (Ø), (ś)	v, k	y (Ø), v	h, Ø, (ḍ)
Reflexes are listed in their apparent order of frequency. Round brackets () indicate reflexes which occur in limited numbers or are atypical. A question mark (?) indicates cases where the text indicates or assumes that the typical reflex is retention, but no examples are listed in the overview.							

Some of the examples given in the source texts are listed in the following tabulation:

Source		Sanskrit > Gāndhārī
Aś Ins.	ś	ś > ś: -daśa- > -daśa- ś > s: anuśocana- > anusocana-
	ṣ	ṣ > ṣ: śuśrūṣa- > suśruṣa- ṣ > ś: manuṣa- > manuśa- ṣ > s: abhiṣikta- > abhisita-
	s	s > ś: śāsana- > śaśana- (assimilation)
	h	h > h: iha > iha, aham > aham h > Ø: iha > ia, āha > aa (M)

BLS1-1	ś	ś > ś: <i>cāturdiśe</i> > <i>cadudiśami</i> ś > h: <i>yoniśaḥ</i> > <i>yoṇiho</i> ś > Ø: <i>vaśībhūtaḥ</i> > <i>baīhodu</i> , <i>*vaṅkiśaḥ</i> > <i>baki[a]</i>
	ṣ	ṣ > ṣ: <i>trṣitam</i> > <i>tiṣidu</i> , <i>sugatiṣu</i> > <i>sughadiṣu</i>
	s	s > s: <i>-saha-</i> > <i>-saha-</i> , <i>prāsādikāḥ</i> > <i>prasadio</i>
	h	h > h: <i>-saha-</i> > <i>-saha-</i> , <i>lohitam</i> > <i>luhitu</i> h > Ø: <i>*bhuñjahi</i> > <i>bhujai</i>
BLS1-2	ś	ś > ś: <i>prakāśet</i> > <i>praghaśe</i> , <i>saṃprakāśitaḥ</i> > <i>[s](*a)praghaśi[d]e</i> , <i>akuśalāḥ</i> > <i>akuśala</i>
	s	s > s: P <i>anāthapiṇḍikassa</i> > <i>aṇasapiḍiasa</i> , <i>upasaṃkramitvā</i> > <i>uasakramita</i>
	h	h > h: <i>saha-</i> > <i>saha-</i> h > s: <i>prahāṇa</i> > <i>[pra]saṇa</i> (possibly developed under the influence of <i>pradhāṇa</i> > <i>pra[sa]ṇa</i>)
BLS1-3	s	s > s: P <i>bodhisattva-</i> > <i>bohisatva-</i> s > h: P <i>paccasārī</i> > <i>pracahari</i> s > h > Ø: P <i>paccayāse</i> > <i>*pracayahe</i> > <i>praceae/i</i>
	h	h > h: P <i>jahāti</i> > <i>jahadi</i> , P <i>pajahāti</i> > <i>prahadi</i> ; h > Ø: P <i>saroruham</i> > <i>saleloa</i>
BLS2-1	ś	ś > ś: <i>prādeśikāḥ</i> > <i>pradeśige</i>
	ṣ	ṣ > ś: <i>ābhāṣate</i> > <i>abhaṣadi</i>
	s	s > s: <i>-kasam-</i> > <i>-g[a]sa-</i> s > h: <i>-kasam-</i> > <i>-[geha]-</i> (possibly also Skt. <i>geha-</i>)
	h	h > h: <i>anugraha-</i> > <i>anugraha-</i> , <i>aham</i> > <i>aha</i> / <i>aho</i> , <i>udāharanena</i> > <i>[u]dahara[ne](*no)</i>
BLS2-2	ś	ś > ś: <i>praviśati</i> > <i>praviśadi</i> , ś > Ø: <i>takṣaśilārājah</i> > <i>takṣaīlara[ya]</i>
	s	s > s: P <i>bodhisattva-</i> > <i>bosisatva-</i> , <i>corasamākulaḥ</i> > <i>corasamagule</i>
	h	h > h: <i>grhītāḥ</i> > <i>grahida</i> h > Ø: <i>grhītāḥ</i> > <i>ghini[do]</i> (aspiration shift)
BLS4-1	ś	ś > ś: <i>*ākāśa-</i> > <i>agaśa-</i> , <i>aśukla-</i> > <i>aśukra-</i> , <i>kleśa-</i> > <i>kileśa-</i>
	ṣ	ṣ > ṣ: <i>skandheṣu</i> > <i>kadheṣu</i> , <i>*dveṣaḥ</i> > <i>doṣa</i> , <i>puruṣaḥ</i> > <i>puruṣa</i>
	s	s > s: <i>cetasah</i> > <i>cedasa</i> , <i>-nivāsā-</i> > <i>-ṇivasa-</i> , <i>manasā</i> > <i>maṇaso</i> , <i>vasana-</i> > <i>vasaṇa-</i>
	h	h > h: <i>anupahata-</i> > <i>anuahada-</i> , <i>aspr̥hāluḥ</i> > <i>aprihalu-</i> , <i>prahāṇam</i> > <i>prahaṇo</i> , <i>moha-</i> > <i>moha-</i> h > ś: <i>iha</i> > <i>iśa</i>

BLS9-1	ś	ś > ś: <i>avaśāṭayitvā</i> > <i>ośaḍaita</i> , <i>praśastam</i> > <i>praśasta</i> (Salomon 2000: 223)
	ṣ	ṣ > ṣ: <i>-viṣāṇakalpo</i> > <i>-viṣaṇagapo</i> , <i>-viṣūkāni</i> > <i>-viṣoṇi</i> , <i>kāṣāya-</i> > <i>kaṣa(*ya)-</i> (p. 229)
	s	s > s: <i>anartha-sevinam</i> > <i>aṇartha-sevi</i> (p. 227) s > ṣ: <i>pipāsam</i> > <i>pivaṣa</i>
	h	<i>avihesaka-</i> > <i>avihesao-</i> , <i>parigrahān</i> > <i>parigrahe</i> , <i>sahāyam</i> > <i>sahaya</i>
BCS5-1	ś	ś > ś: <i>akulaśa-</i> > <i>akuśala-</i> , <i>kulaśa-</i> > <i>kuśala-</i> , <i>deśa-</i> > <i>deśa-</i> , <i>nāśa-</i> > <i>ṇaśa-</i> ś > ś: <i>anuśaṃsa-</i> > <i>aṇuśaśa-</i> , <i>aśubha-</i> > <i>aśuha-</i> (after prefixes)
	ṣ	ṣ > ṣ: <i>asatpuruṣa-</i> > <i>aṣapuruṣa-</i> (Schlosser 2016: 296)
	s	s > ṣ: <i>āsaptika-</i> > <i>aṣatia</i> , <i>asatpuruṣa-</i> > <i>aṣapuruṣa-</i> (p. 296), <i>caitasika-</i> > <i>cedaṣia-</i>
	h	h > ś: <i>iha</i> > <i>iṣemi</i> h > ṣ: <i>prahāṇa</i> > <i>praṣaṇa</i>
RS-5	ś	ś > ś: typically stable, e.g., <i>aśubhā</i> > <i>aśua</i> ś > ṣ / s: <i>aśuci</i> > <i>aṣuyi</i> / [a]suyi[a]
	ṣ	ṣ > ṣ: typically stable, e.g., <i>purīṣam</i> > <i>puriṣa</i> ṣ > ṣ: <i>kulmāṣaḥ</i> > <i>k[u]ma[ṣa]</i>
	s	Typically, s > ṣ: e.g., <i>vadasi</i> > <i>vaḍeṣi</i>
	h	h > h: typically stable, e.g., <i>sahagata</i> > <i>sahagada</i> h > Ø: <i>plīhakaḥ</i> > <i>priao</i> , <i>prajahatha</i> > <i>pajaaṣa</i> h > ś: <i>iha</i> > <i>iṣe</i>
RS-12	ś	ś > ś: typically stable, e.g., <i>viśeṣa-</i> > <i>viṣeṣa-</i> , <i>sarvaśaḥ</i> > <i>sarvaśa</i> , <i>pradeśam</i> > <i>pradeśi</i> , <i>ākāśaḥ</i> > [aga]śa ś > h: <i>pratiśāmayati</i> > [padi]hamedī ś > s: <i>praviśa</i> > (*pra)[vi]s[*a]
	ṣ	Typically, ṣ > ṣ: <i>aśrauṣīt</i> > [a]śroṣi, <i>tuṣitāḥ</i> > <i>tuṣida</i> , <i>nyaṣīdat</i> > <i>ṇiṣīdi</i> , <i>viśeṣa-</i> > <i>viṣeṣa-</i> ; <i>samādāpayeti</i> > <i>samadaveṣi</i> appears to be a scribal slip.
	s	s > s: typically stable, e.g., <i>upasāmpadya</i> > <i>vasapaca</i> , <i>nivāsam</i> > <i>ṇivaso</i> , <i>prītisukham</i> > <i>pidisuho</i> s > ṣ: <i>upasāmpadya</i> > <i>vaṣapaja</i> , BHS <i>upasāmkramitvā</i> > <i>uaṣakramita</i> , <i>cetasā</i> > <i>cedaṣa</i> .
	h	h > h: <i>muhūrtam</i> > <i>mahuta</i> , <i>-hara-</i> > <i>-hara-</i> many times, e.g., <i>viharanti</i> > <i>viharati</i> . h > ś: <i>iha</i> > <i>iṣe</i> h > Ø: <i>-ehi</i> > <i>-ehi</i> / <i>-ei</i> / <i>-e</i>
RS-14	ś	ś > ś: <i>*trītrimśam</i> > [tri]triṣe

	ś	ś > ś: <i>maharṣeḥ</i> > <i>mahirṣ[i]ṇa</i> (refer to the text discussion below regarding these consonant clusters)
	s	s > s: <i>manasā</i> > <i>maṇasa</i>
	h	h > v / ḳ: <i>aḥam</i> > <i>avū</i> , <i>saptāḥam</i> > <i>sataḥkam</i> (unusual developments)
RS-19	h	h > y: <i>gr̥ha-</i> > <i>giyi-</i> h > v: <i>aḥam</i> > <i>avū</i>
RS-20	ś	ś > ś: typically stable, e.g., <i>bhāṣitam</i> > <i>bhaṣiḍe</i>
	s	s > s: <i>upāsakam</i> > <i>ausao</i> s > ṣ: <i>P *abhisankharonti</i> > <i>aviṣakhareti</i>
	h	h > h: <i>P imehi</i> > <i>imehi</i> h > Ø: <i>-paridāḥam</i> > <i>-paraḍaa</i> h > ḍ: <i>-hantatarah</i> > <i>-ḍaḍaro</i> , may be due to orthographic similarity.

The typical reflexes indicate that the intervocalic single sibilants are preserved in Gāndhārī, including in the Aśokan inscriptions, and the distinction between the three types of sibilants is maintained, compared to Pali for example where these are merged (Oberlies 2019: 114-115). The distinction between the sibilants is also evident in reflexes of consonant clusters involving sibilants (e.g., §§ 8.3.1.3, 8.3.3, 8.3.4.4.3), as shown in the RS-14 examples listed in the above table. While the retained sibilants are the most commonly recorded forms, some alternate reflexes are also recorded, which include lenition to *h*, and the use of diacritic marked reflexes by the BCS5 and RS scribes. While the glottal fricative *h* appears to be preserved in many instances, variable reflexes which appear to indicate some form of lenition or deletion are also recorded by most scribes.

While the recording of unchanged ś is the main reflex for the palatal sibilant, the use of the alternate diacritic marked ṣ by the BC and RS scribes has been interpreted to indicate that the sibilant was voiced as [j], rather than remaining unvoiced as [ç]²²⁹. This form has also been used and by the BL scribe 14 (Baums 2009: 150)²³⁰. Alternate reflexes such of *h* and Ø which suggest a form of lenition are uncommon and assumed to result from scribal laxity or confusion, although these sometimes arise in a palatal environment, such as in Skt. *yonīśaḥ* > G *yoṇiḥo* and **vaṇkiśaḥ* > *baki[a]* for example. The retroflex sibilant ṣ is recorded unchanged in most cases, although the use of diacritic marked ṣ by the BC scribe, and in the RS-12 manuscript, appears to represent an attempt by these scribes to convey the allophonic voiced pronunciation [z] of this phoneme in intervocalic

²²⁹ Baums 2009: 150; Silverlock 2015: 244.

²³⁰ This scribes work is not included in the selected source materials.

positions²³¹. The dental sibilant *s* is also recorded unchanged in most cases, although use of the diacritic marked *ṣ* by the BC and RS scribes is assumed to represent the voiced pronunciation [z], and the widespread use of both forms suggests they may have become free graphic variants, both representing the same sound²³². The intervocalic voicing of various sounds, including stops and fricatives, is very common in the world's languages (Campbell 2016: 35), and this interpretation is quite reasonable. The reflexes of the aspirated dental stops (*th*, *dh*) as (*s*, *ṣ*) (§ 8.2.1.2.4) are assumed to represent lenition as the voiced dental or alveolar voiced fricatives [ð] or [z]. This raises questions about the potential merger of these sounds with the sibilant *s*, and the subsequent development of homonym pairs (§§ 8.2.1.2.4, 8.2.1.2.6.2).

Intervocalic *h* is unchanged in most cases, although deletion is also recorded occasionally. While it is assumed that *h* represents a voiced glottal fricative (Baums 2009: 149), it has been suggested that it may have become unvoiced, creating a hiatus between the vowels (Silverlock 2015: 246). The development of *iha* > *iśa* / *iṣe* is an anomalous but common Gāndhārī reflex, likely to have resulted from the palatal environment, also reflected in the development of *a* > *e*. The unusual reflexes of *v* and *k* may reflect a common lenition and levelling of their pronunciation with *h* (Salomon: 2008: 340).

Summary: an overview of the Gāndhārī reflexes of the OIA intervocalic fricatives is presented in Table 8.2.3.2b.

Table 8.2.3.2b: Overview of intervocalic fricative reflexes in Gāndhārī

OIA Fricative	Suggested G phoneme	Assumed sound (Sibilants become voiced)	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-ś-/	/-ś-/	[j]	/ś/ ≡ {/ś ^h /, /ś ^{alt} /, (/Ø/, /h/, /ś/, /ṣ/, /h/) ^{alt} }
/-ṣ-/	/-ṣ-/	[z]	/ṣ/ ≡ {/ṣ ^h /, /ṣ ^{alt} /, (/ṣ/) ^{alt} }
/-s-/	/-s-/	[z]	/s/ ≡ {/s ^h /, /s ^{alt} /, (/Ø/, /h/) ^{alt} }
/-h-/	/-h-/	[ɦ]	/h/ ≡ {/h ^h /, (/Ø/, /ś/, /ṣ/) ^{alt} }

²³¹ Baums 2009: 150; Silverlock 2015: 244-245.

²³² Baums 2009: 150; Silverlock 2015: 245-246; Marino 2017: 126.

8.2.3.3 Word final

In OIA sandhi rules word final sibilants are disallowed and are replaced by other consonants or visarga *-ḥ* in some instances. As noted in § 8.2.1.3, word final consonants are lost in Gāndhārī, which also applies to visarga (e.g., Baums 2009: 212-213).

8.2.4 Semi-vowels

8.2.4.1 Word initial

OIA word initial semi-vowels generally remain stable in Gāndhārī²³³. A small number of atypical developments have been observed, including:

- Initial *v* is sometimes represented as *b*, as in *vaśībhūtaḥ* > *baīhodu*, *varṇavān* > *baṇave* for example (Salomon 2008: 116), which is also typical in some versions of Pali. The alternation of *v* and *b* is also seen in some *b* reflexes, as in *ucyate* > *bucadi*, *vucadi* for example, which is not uncommon in Gāndhārī (Allon 2001: 78).
- Initial *v* has also been represented by *d-* and *bh-*, as in Skt. *vihaṅgamaḥ* > G *dihaghamsa* and P *vinaḷīkatā* > *bhiriḍi-ghama* for example. The writing of *di-* for *vi-* may be the result of scribal error, while the use of aspirated *bh-* appears problematic (Allon 2001: 78).
- Metathesis of original *v-* in Skt. *vr̥kṣa-* > G *rukṣa-* / *rokṣa-*, which may also develop as *vr̥* > **vru* > *ru* / *ro*²³⁴.

Summary: OIA word initial semi-vowels generally remain stable in Gāndhārī, and the development from Skt. > G is typically represented as /R_x-/ > /R_x-/, where x is representative of the same place feature.

8.2.4.2 Word medial

The Gāndhārī reflexes of OIA single intervocalic semi-vowels recorded in the Aśokan inscriptions and Gandhāran Buddhist manuscript sources are presented in Table 8.2.4.2a.

²³³ E.g., Mehendale 1948: 13-15, 215-216, 301; Allon 2001: 78-79; Glass 2007: 114, 117; Marino 2017: 117.

²³⁴ Allon 2001: 77; Glass 2007: 117.

Table 8.2.4.2a: Gāndhārī reflexes of OIA intervocalic semi-vowels – all sources

Era	3 BCE	1 CE					
Source:	Aś M	BLS1-1	BLS1-2	BLS1-3	BLS2-1	BLS2-2	BLS4-1
pages:	13-15, 215-216	108, 116	85-86	40-43	128-130	25-27	137, 148-149
OIA -y-	<i>Ø, y, j</i>	<i>Ø, y</i>	<i>Ø, y</i>	<i>Ø</i>	<i>y, (g), (j), (Ø)</i>	<i>g, Ø</i>	<i>Ø, y, g</i>
-r-	<i>r</i>	<i>r</i>	<i>r</i>	<i>r, (l), (d)</i>	<i>r</i>	<i>r</i>	<i>r, (l)</i>
-l-	<i>l (ḍ)</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>
-v-	<i>v</i>	<i>v, (Ø)</i>	<i>v, (m)</i>	<i>v</i>	<i>(?)</i>	<i>v, (Ø)</i>	<i>v</i>
Era	1 CE	2 CE					
Source:	BLS9-1	BCS5-1	RS-5	RS-12	RS-14	RS-19	RS-20
pages:	81-86	76-77	114, 117-118	229, 240-243	338-340	55	117, 125
-y-	<i>Ø, y</i>	<i>Ø, y</i>	<i>Ø, y</i>	<i>Ø, y, (h)</i>	<i>(?), (Ø)</i>	<i>Ø</i>	<i>y, Ø</i>
-r-	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>		<i>r, (ḍ)</i>
-l-	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>		<i>l, (ḍ)</i>
-v-	<i>v</i>	<i>v</i>	<i>v, (Ø)</i>	<i>v</i>	<i>v</i>		<i>v</i>
Reflexes are listed in their apparent order of frequency. Round brackets () indicate reflexes which occur in limited numbers or are atypical.							

The alternative forms of the Gāndhārī reflexes recorded in the source materials are summarised in the following formulae.

The palatal approximant *y*:

Aś Ins: $y > \{y, \emptyset, (j)\} / V_V$

BL, RS and BCS5 Scribes: $y > \{\emptyset, y, (g)\} / V_V$

The alveolar tap *r*:

Aś Ins: $r > \{r\} / V_V$

BL, RS and BCS5 Scribes: $r > \{r, (l), (ḍ)\} / V_V$

The dental approximant *l*:

Aś Ins: $l > \{l, (ḍ)\} / V_V$

BL, RS and BCS5 Scribes: $l > \{l, (ḍ)\} / V_V$

The labial approximant v :

Aś Ins: $v > \{v\} / V_V$

BL, RS and BCS5 Scribes: $v > \{v, (Ø)\} / V_V$

Selected examples from the source texts are listed in the following tabulation:

Source		Sanskrit > Gāndhārī
Aś Ins	y	$y > Ø$: *ekatiya- > ekatia- $y > y$: priya- > piya- $y > j$: mayūra- > majura- (one instance)
	r	$r > r$: caraṇa- > -caraṇa-, cira- > cira-
	l	$l > ḍ$: Keralā- > Keraḍa-
	v	$v > v$: eva > eva
BLS1-1	y	$y > Ø$: udariyakam > udario, cāmpayaḥ > capeo, priyaḥ > prio, rajanīyāni > rayāṇia, hṛdayam > hideo $y > y$: ayam > aya, janakāyaḥ > janaghayu, niṣkaṣayaḥ > nikaṣaya, bhūyaḥ > bhuyo
	r	$r > r$: udariyakam > udario, nirayeṣu > ṇiraghehi, vārāṇasīpure > baraṇasīpure
	l	$l > l$: brahmaloka- > bramal«u»gha-
	v	$v > v$: bhavaḥ > bhavu $v > Ø$: evam > ea appears to be scribal error. In clusters (-vr-, -rv-) > -b- (Refer to § 8.3.4.2.4, § 8.3.4.4.4)
BLS1-2	y	$y > Ø$: -indriyam > -idhrio, samayena > samaeṇa, BHS vipūyakam > v[i]puao $y > y$: P samayam > sama[ya], kāyena > kayeṇa, P kāyindriyam > kayidhri
	r	$r > r$: adhikaraṇam > asiarāṇam (Allon 2001: 80)
	l	$l > l$: śīlavantam > śīlamada (p. 85)
	v	Typically stable, but $v > Ø$ in evam > ema
BLS1-3	y	$y > Ø$: dalayitvā > (*da)laīta, P paccayāse > praceae
	r	Typically stable, but $r > d$ in P sārām > sada, $r > l$ in P saroruhām > saleloa

	<i>l</i>	<i>l</i> > <i>l</i> : <i>dalayitvā</i> > (<i>*da</i>) <i>laīta</i> ,
	<i>v</i>	Typically stable
BLS2-1	<i>y</i>	<i>y</i> > <i>y</i> : <i>utpādayati</i> > <i>upadayati</i> , <i>prayate</i> > <i>p(*ra)ya[d](*)mi</i> , <i>pūrvayogaḥ</i> > <i>provayoge</i> ; <i>y</i> > \emptyset : <i>mantrayati</i> > <i>matradi</i> <i>y</i> > <i>g</i> : <i>samvṛtendriyaḥ</i> > <i>sabrudidriḡo</i> , <i>kāranīyam</i> > <i>[ka]ranigo</i> <i>y</i> > <i>j</i> : <i>ājānīyam</i> > <i>ayaneja</i> , <i>ārāgayeyam</i> > <i>aragajo</i> (optative endings)
	<i>r</i>	<i>r</i> > <i>r</i> : <i>nagare</i> > <i>nagare</i> , <i>anyataraḥ</i> > <i>añadaro</i>
	<i>l</i>	<i>l</i> > <i>l</i> : <i>kālagataḥ</i> > <i>kalagado</i>
BLS2-2	<i>y</i>	<i>y</i> > <i>g</i> : <i>indriyānām</i> > <i>idriḡano</i> , <i>dvitīyaḥ</i> > <i>bidiga</i> , <i>vīryendriyaḥ</i> > <i>viridriḡo</i> , <i>pālendriyaḥ</i> > <i>palaīdrago</i> <i>y</i> > \emptyset : <i>mantrayati</i> > <i>matredi</i> (common)
	<i>r</i>	<i>r</i> > <i>r</i> : <i>aparaḥ</i> > <i>avare</i> , <i>parataḥ</i> > <i>parado</i>
	<i>l</i>	<i>l</i> > <i>l</i> : <i>pālendriyaḥ</i> > <i>palaīdrago</i>
	<i>v</i>	<i>v</i> > <i>v</i> : <i>ājivika</i> > <i>ayiviga</i> , <i>sthavikāḥ</i> > <i>thavigo</i> , <i>praviśati</i> > <i>praviśadi</i> , <i>bhavati</i> > <i>hovadi</i> <i>v</i> > \emptyset : <i>yāvat</i> > <i>yaa</i>
BLS4-1	<i>y</i>	<i>y</i> > \emptyset : <i>*āyatana-</i> > <i>aīdaṇa</i> , <i>anunaya-</i> > <i>aṇuṇea-</i> , <i>*udaya-</i> > <i>udea</i> , <i>*kśīyante</i> > <i>kṣiati</i> , <i>prahāya</i> > <i>prahae</i> , among many others. Elision in presence of palatal vowels or following palatalisation of vowels. <i>y</i> > <i>y</i> : <i>*āyatana-</i> > <i>ayadana-</i> , <i>kāya-</i> > <i>kaya-</i> , <i>kṣaya-</i> > <i>kṣaya-</i> , <i>samudaya-</i> > <i>samudaya-</i> <i>y</i> > <i>g</i> : <i>a-neyaḥ</i> > <i>aṇego</i> , <i>āśaya-</i> > <i>aśaga-</i> , <i>samśaya-</i> > <i>śaśaga-</i>
	<i>r</i>	<i>r</i> > <i>r</i> : <i>ārya-</i> > <i>aria-</i> , <i>avasirati</i> > <i>osiradi</i> , <i>puruṣa-</i> > <i>puruṣa</i> , <i>maruṇa-</i> > <i>maruṇa-</i> <i>r</i> > <i>l</i> : once in <i>parīkṣita</i> > <i>palikhaīda</i> , suggested as a borrowing.
	<i>l</i>	<i>l</i> > <i>l</i> : <i>*asprḥālu-</i> > <i>aprihalu-</i> , <i>*mūla-</i> > <i>mula-</i> , <i>*sakala-</i> > <i>sakala-</i>
	<i>v</i>	<i>v</i> > <i>v</i> : <i>anavasruta-</i> > <i>aṇavaṣuda-</i> , <i>anāsrava-</i> > <i>aṇasava-</i> , <i>avidya-</i> > <i>avija-</i> , <i>bhava-</i> > <i>bhava-</i>
BLS9-1	<i>y</i>	<i>y</i> > \emptyset : <i>guptendriyaḥ</i> > <i>gutimdrīo</i> , <i>dvitīyena</i> > <i>dutieṇa</i> , <i>sāmayikam</i> > <i>samaia</i> . Typically elided or left unwritten when vowels on either side dissimilar. <i>y</i> > <i>y</i> : <i>-kāṣāya-</i> > <i>-kaṣaya-</i> , <i>sāhaya-</i> > <i>sahaya-</i>
	<i>r</i>	<i>r</i> > <i>r</i> : P <i>parissayāni</i> > <i>pariśeāni</i>
	<i>l</i>	<i>l</i> > <i>l</i> : <i>pratīlabdha-</i> > <i>paḍila[dha]-</i> , <i>vācābhilāpo</i> > <i>vyabh.lavo</i>
	<i>v</i>	<i>v</i> > <i>v</i> : P <i>apātivatto</i> > <i>uvativuto</i>
BCS5-1	<i>y</i>	<i>y</i> > \emptyset / <i>y</i> : Commonly elided in combination with vowels, but also retained in some cases: <i>bhūya-</i> > <i>bhuya-</i> , <i>svaya-</i> > <i>svaya-</i> , <i>śriya-</i> > <i>śriya-</i> .

	<i>r</i>	<i>r</i> > <i>r</i> : <i>apari-</i> > <i>apari-</i> , <i>apara-</i> > <i>apara-</i>
	<i>l</i>	<i>l</i> > <i>l</i> : <i>palāla-</i> > <i>palala-</i>
	<i>v</i>	<i>v</i> > <i>v</i> : <i>bhāvanā</i> > <i>bhavaṇa</i>
RS-5	<i>y</i>	<i>y</i> > Ø: Typically elided between dissimilar vowels, e.g., <i>antardhāyīṣyāmi</i> > <i>anta[ra]-ṣaiśama</i> <i>y</i> > <i>y</i> : Mostly retained between similar vowels, e.g., <i>kāyam</i> > <i>kaya</i>
	<i>r</i>	<i>r</i> > <i>r</i> : <i>katarah</i> > <i>kaḍara</i>
	<i>l</i>	<i>l</i> > <i>l</i> : <i>pādataḷāt</i> > <i>paḍatala</i>
	<i>v</i>	<i>v</i> > <i>v</i> : <i>eva</i> > <i>eva</i> <i>v</i> > Ø: possible but uncertain due to damaged manuscript.
RS-12	<i>y</i>	<i>y</i> > Ø: Typically elided between dissimilar vowels, e.g., <i>āyusman</i> > <i>aīspa</i> , <i>anusmareyuh</i> > <i>aṇusparea</i> , <i>kṣamaṇīyam</i> > <i>khamañio</i> , <i>trayaḥ</i> > <i>trae</i> . <i>y</i> > <i>y</i> : Often retained between similar vowels, e.g., <i>kāyakarman</i> > <i>kayīame</i> , <i>trayaḥ</i> > <i>trayo</i> , <i>vayam</i> > <i>vaya</i> .
	<i>r</i>	<i>r</i> > <i>r</i> : <i>uttarim</i> > <i>utvara</i> , <i>dūre</i> > <i>[dur]e</i> , <i>parikṣiṇa</i> > <i>parakṣiṇa</i> , <i>avicāram</i> > <i>avicara</i> , among many examples.
	<i>l</i>	<i>l</i> > <i>l</i> : <i>sulabdhalābhāḥ</i> > <i>suladhalava</i> , <i>akuśalaiḥ</i> > <i>akuśal[e](*hi)</i> <i>l</i> > <i>r</i> : <i>P kimbilo</i> > <i>kibhire</i> (One of the main characters - possibly an eastern influence?)
	<i>v</i>	<i>v</i> > <i>v</i> : <i>ava-</i> > <i>ava-</i> , <i>avi</i> > <i>avi-</i> , <i>evam</i> > <i>eva</i> , <i>bhagavantam-</i> > <i>bhagavata</i> , <i>sadevakaḥ</i> > <i>sadevao</i> , among many others.
RS-14	<i>y</i>	<i>y</i> > Ø: One uncertain case of elision.
	<i>r</i>	<i>r</i> > <i>r</i> : <i>ācaritam</i> > <i>-(a)carido</i>
	<i>l</i>	<i>l</i> > <i>l</i> : <i>akhila-</i> > <i>agila-</i> , <i>śīlavataḥ</i> > <i>(*śila)vade</i> ,
	<i>v</i>	<i>v</i> > <i>v</i> : <i>śrāvakaḥ</i> > <i>śavaka</i>
RS-19	<i>y</i>	<i>y</i> > Ø: <i>-nikāyam</i> > <i>-niao</i> , BHS <i>-paryāyam</i> > <i>-paīao</i> , <i>karaṇīyam</i> > <i>karaṇia</i> , <i>samayena</i> > <i>samaeṇa</i> .
RS-20	<i>y</i>	<i>y</i> > <i>y</i> : Typically stable, e.g., <i>kāyāḥ</i> > <i>kaya</i> <i>y</i> > Ø: <i>samayam</i> > <i>samae</i>
	<i>r</i>	<i>r</i> > <i>r</i> : <i>anyataraḥ</i> > <i>aṇearo</i> . Metathesis one case: <i>audārikāḥ</i> > <i>oraḍi</i>
	<i>l</i>	<i>l</i> > <i>l</i> : <i>-jvalitāḥ</i> > <i>-jalida</i> . Metathesis one case: <i>biḍala-</i> > <i>bilaḍa-</i>
	<i>v</i>	<i>v</i> > <i>v</i> : <i>śrāvastī-</i> > <i>śavasti-</i>

The typical reflexes indicate that the intervocalic OIA semi-vowels *y*, *r*, *l* and *v* are generally preserved in Gāndhārī, including in the Aśokan inscriptions, although some variations occur in the record. Most commonly, *y* is often not recorded between dissimilar vowels, particularly when these include at least one of the high or high-mid front vowels

such as *i* and *e*, while it is typically retained when it occurs between two *a*-vowels, as in *aya*²³⁵. This development is also recorded in the Aśokan inscriptions, seen in Skt. **ekatiya-* > G *ekatia-* for example, although *y* is also preserved in *priya-* > *piya-*. In these cases, *y* possibly emerges as a weakly pronounced *ya-śruti* (Misra 1967: 166) as the tongue moves closer to the palate for at least one of the vowels, as in *a^ye*, *a^yi*, *i^ya*, *i^yo* and *e^yi* for example. The scribes apparently did not consider it necessary to record this sound and it is mostly “left unwritten”. The recording of *y* > *g* by the BL scribes 2 and 4 appears to represent a historical spelling (Baums 2009: 148), which is only seen on an occasional basis, while the recording of *j* in optative endings by BL scribe 2 may be a MIA borrowing (Lenz 2003: 130). In another irregularity, both *r* and *l* are recorded as *ḍ* in the RS-20 manuscript, as in Skt. *biḍāla-* > G *bilaḍa-* for example, although these are single cases involving metathesis rather than a real phonetic change, while the recording of *Kerala-* > *Keraḍa-* in the Aśokan inscriptions may be due to graphic similarities between the graphemes, or alternately could provide evidence for the presence of the phoneme *ḷ* in OIA. The writing of *mayūra-* > *majura-* in the Aśokan inscriptions suggests an inverse spelling reflecting the merger of *j* with *y* (§ 8.2.1.2.2). The two recorded cases where *v* is elided may be the result scribal error (Salomon 2008: 116).

Summary: an overview of the Gāndhārī reflexes of the OIA intervocalic semi-vowels is presented in Table 8.2.4.2b.

Table 8.2.4.2b: Overview of intervocalic semi-vowel reflexes in Gāndhārī

OIA Semi-vowel	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-y-/	/-y-/ /-Ø-/	[j] [j̥]	/y/ ≡ {/y/ ^h , /Ø/ ^{alt} , (/g/, /j/) ^h } Elision or <i>ya-śruti</i> when at least one high-mid or high vowel.
/-r-/	/-r-/	[r]	/r/ ≡ /r/ ^h
/-l-/	/-l-/	[l]	/l/ ≡ /l/ ^h
/-v-/	/-v-/	[v]	/v/ ≡ /v/ ^h

²³⁵ Salomon 2000: 86; Allon 2001: 85; Baums 2009: 148; Silverlock 2015: 240.

8.2.4.3 Word final

The semi-vowels *y*, *l*, and *v* do not occur word finally in OIA, while *-r* is replaced by visarga (Macdonell 1927: 13-14).

8.2.5 Scribal variations

A number of variations are observed in the phonological reflexes recorded by different scribes, including in the inconsistent use of diacritic markers in the recording of sound changes, and recording of the orthographic merger of the dental and retroflex nasals using alternate graphemes. At the same time, other significant phonological developments such as the elision of some consonants in the environment of palatal vowels appear to have been recorded reasonably consistently by all scribes.

8.2.5.1 Inclusion of diacritic markers to distinguish sound changes (§ 8.2.1.2.6.1)

This innovation appears to have been predominantly used by the RS scribe²³⁶, mostly in the reflexes of the velar and dental unaspirated stops, as in *k / g > ḡ* and *t / d > ḍ*. This is also seen in reflexes of the aspirated dental stops, as in *th / dh > ṣ*, and appears to have been used in order to distinguish this reflex from the dental sibilant *s*. These diacritics were also used, although to a more limited extent, by the BC scribe 5, while they are typically not used by the BL scribes, apart from the use of *ṣ* by the BL scribe 2. These innovations suggest that RS scribe adopted a more nuanced approach in recording the changing or variable sounds of the language than the BL scribes. At the same time these reflexes are not used consistently, either through hesitancy or laxity, seen in Skt. *loke > G loge / loga* (RS-5), *loke > loga / logo* (RS-20), *bhagavant > bhagava / bhagava / bhayava* (RS-12), *śrāvaka- > śavaḡa- / śavaka-* (RS-14), and *yuśmākam > tuspahu* (RS-5) compared with *upāsakam > uasao* (RS-20) for example. However, the use of underscore diacritics in *t / d > ḍ* appears to be applied more consistently by the scribe than with *k / g > ḡ*, although some variations such as *utpadyate > apacadi / apacati* (RS-20) are recorded. The development of *th / dh > ṣ* is also recorded inconsistently, as in the variations of *yathā > yaṣa / yasa* seen in most of the RS manuscripts.

²³⁶ This also appears to be a feature in those manuscripts of the IM Collection which have been studied to date (M. Allon, personal communication, February 27, 2023).

8.2.5.2 Variable use of *n/ṇ* (§ 8.2.2.2.2)

The Gandhāran Buddhist scribes appear to have recognised that there was little if any phonemic distinction between the dental and retroflex nasals *n* and *ṇ* in Gāndhārī, and this is observed in the use of orthographic *ṇ* to represent both phones by most of the BL scribes as well as the BC and RS scribes. The use of orthographic *n* was however adopted by the BL scribe 2, and it is suggested that this is the more appropriate form of phonemic representation for these phones, pointing towards a more considered understanding of the phonetics by this scribe. At the same time the RS scribe appears to have followed the more commonly accepted orthographic practice in this regard.

8.2.5.3 Elision of unaspirated velar and palatal stops in a palatal environment

The elision (Ø) of Sanskrit intervocalic unaspirated velar and palatal stops in the environment of palatal vowels appears to have been a widespread practice, which was followed with reasonable consistency by all of the Gāndhārī scribes. Some variations are observed in most of the manuscripts, however, where the stops are either not elided in palatal environments or elided in non-palatal environments. Some of these patterns are shown in the following examples. For the velar stops: Skt. *anadhika-* > G *anasia-*, *pāpikām* > *pavia*, but BHS *antraguṇam* > (**a*)*traūṇa* (BLS1-1); *bhaginī* > *bhaiṇi* (BLS1-2); *pārāyaṇikaḥ* > *pa[r]ayaṇio*, but *kṛtakah* > *kridaga*, *prayoga-* > *prayoa-* (BLS4-1); *sāmayikam* > *samaia* (BLS9-1), *caitasika-* > *cedaṣia-*, *paligodhena* > [*a*]*palioṣeṇa*, *kapālakena* > *kavalaeṇa*, but *-bhogin* > *-bhoyi*, *suvaīrāga-* > *suverao-* (BCS5-1); *abhyavakāśa* > *abhoāśa* but *pratīkūla-* > *paḍikula* (RS-5); *nandikaḥ* > *ṇadio* but *kāya-karman-* > *kayīame-* (RS-12); *-antike* > *-ṣatia*, *-nikāyam* > *-ṇiao* (RS-19); **mṛga* > *mia*, but *upāsakam* > *uasao*, and *ayo-guḍaḥ* > *aya-ūḍa* (RS-20). For the palatal stops: Skt. **paricaritvā* > G *pariarita*, *abhijānāmi* > *abhiaṇami*, (BLS1-1); *avaca* > *a[va]i*, *vicaya* > *vie* (BLS1-2); P *vanathajā* > *baṇahea[e]* (BLS1-3); *brahmacaryiṣaṇā* > *bra*(**maīyeṣana*), but *niyojanam* > *ṇiyoaṇa* (BLS4-1), which may be influenced by the preceding *-iy-*; *ācakṣante* > [*a*]*ikṣati* (RS-12), and *sacet* > *sai* (RS-14).

8.2.5.4 Elision of *y* in palatal environments

The elision of intervocalic *y* in palatal environments appears to have been a widespread practice followed by all of the scribes covered by the documents used in this

study, and while some variations are observed these do not appear to be widespread, suggesting a clear awareness by all of the scribes. Examples include: Skt. *udariyakam* > G *udario*, but *hṛdayam* > *hideo* (BLS1-1); *-indriyam* > *-idhrīo*, but *samayena* > *samaeṇa* and *kāyena* > *kayeṇa* (BLS1-2); P *paccayāse* > *praceae* (BLS1-3); **kṣīyante* > *kṣiati*, *anunaya*-> *aṇuṇea*- and *prahāya* > *prahae* (BLS4-1) suggest elision following the palatalisation of vowels; *dvitīyena* > *dutieṇa*, *sāmayikam* > *samaia* (BLS9-1); *kṣamaṇīyam* > *khamāṇio*, *trayaḥ* > *trae* (RS-12); *-nikāyam* > *-niao*, BHS *-paryāyam* > *-paiao* (RS-19), and *samayam* > *samae* (RS-20).

8.2.5.5 Overview

The differences in use of diacritic marked reflexes between the RS and BL scribes indicate a divergence in the graphic habits of these scholarly communities. At the same time, the use of orthographic *ṇ* to represent both *n* and *ṇ* by the RS and BC scribes, and most of the BL scribes, along with the regular elision of the velar and palatal stops and *y* in the environment of palatal vowels by all scribes, indicate that a number of common graphic habits were shared between these scholarly communities. Whatever the origin of these developments, their use was widespread among the Gāndhārī Buddhist communities, whether shared through synchronic or diachronic relationships within and between the communities. It is interesting that the apparent origin of both BL and RS manuscripts was in monasteries near Haḍḍa in Afghanistan, while the BC manuscripts possibly originated from around Bajaur in Pakistan, demonstrating the geographic spread of these graphic habits. And while the BL manuscripts appear to date from the first century, the RS and BC manuscripts possibly date from the second century CE, suggesting a diachronic evolution in the use of diacritics to mark sound variations or changes occurring in the language.

8.2.6 Summary of changes – uncombined single consonants

8.2.6.1 General patterns of change

Although the widespread use of variable reflexes representing historic, transitional and advanced forms confuses understanding of the changes which occurred, a number of general patterns of development for the single consonants have been discerned from the available record. These include:

- Merger of the voiceless stops with their voiced counterparts

- A general tendency for lenition of the intervocalic voiced stops to fricatives or approximants, representing change in progress. Evidence suggests that deletion was not typical, and most stops were subject to phonetic weakening while retaining their phonemic status as stops.
- The OIA nasals *m* and *n* are retained, the distinction between *ṇ* and *n* was lost, and *ṇ* only arises in geminate reflexes of OIA consonant clusters.
- The three OIA sibilants *ś*, *ṣ* and *s* are retained.
- The OIA semi-vowels *y*, *r*, *l* and *v* are typically retained, although *y* tends to be elided in the environment of palatal vowels.

8.2.6.2 Gāndhārī phonological system

A chart of the Gāndhārī phonological system is presented in Table 8.2.6.2. This is based on the chart presented in Table 5.0, and includes representations of both word initial and word medial phonemes. This allows the slightly modified forms suggested for some of the word medial cases to be indicated. Some of the variant forms of reflexes recorded in the source materials for the medial cases are also indicated.

8.2.6.3 Comparison with Torwali phonological system

Some of the more detailed phonological features of the modern day Torwali language, as recorded by Lunsford (2001: 9-26), include the following:

- Both voiceless as well as voiced stops are retained medially, although their pronunciation is weakened, so that for example *g* is typically rendered as the velar fricative [ɣ], *q* and *d* are usually weakened to the flaps [ɾ] and [ɽ], *b* to [w] and *p* to [f].
- Palatal stops appear to have weakened to affricates, recorded as [č] and [j], to the extent that these are classified as affricates in the phonemic inventory.
- There are no aspirated voiced stops, while the aspirated voiceless stops including *kh*, *th*, *th* and *ph* are retained, although these appear to be less common in medial than initial positions.
- There is a class of affricate phonemes, including dental and retroflex affricates along with the palatals.

Table 8.2.6.2: The Gāndhārī Phonological system – suggested representation of word initial and medial single consonants

Consonants		Labial		Dental		Retroflex		Palatal		Velar	
		unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced	unvoiced	voiced
Stops	wi	<i>p-</i>	<i>b-</i>	<i>t-</i>	<i>d-</i>			<i>c-</i>	<i>j-</i>	<i>k-</i>	<i>g-</i>
unaspirated	wm	>	<i>-ḅ-</i>	>	<i>-ḍ-</i>	>	<i>-ḍ-</i>	>	<i>-ḷ-</i>	>	<i>-g-</i>
	≡	{/v/ ^{adv} , /Ø/ ^{alt} , (/b/ ^h)}		{/d/t/ ^h , /ḍ/ ^{tr} , (/Ø/ ^{alt})}		{/ḍ/ ^h }		{/y/ ^{adv} , (/Ø/ ^{adv} , /j/ ^h)}		{/g/ ^h , /gh/ ^{alt} , /ḡ/ ^{tr} , /Ø/ ^{adv} }	
Stops	wi	<i>ph-</i>	<i>bh-</i>	<i>th-</i>	<i>dh-</i>			<i>ch-</i>	<i>jh-</i>	<i>kh-</i>	<i>gh-</i>
aspirated	wm	<i>(-ph-)</i>	<i>-ḅh-</i>	>	<i>-ḍh-</i>	>	<i>-ḍh-</i>		<i>(-jh-)</i>	>	<i>-ḡh-</i>
	≡	{/h/ ^{adv} , /bh/ ^h , /v/ ^{alt} , (/Ø/ ^{alt})}		{/ṣ/ ^{adv} , /s/ ^{alt} , (/dh/ ^h)}		{/ḍh/ ^h }		{/h/ ^{adv} , /g/ ^h , (/Ø/ ^{adv})}			
Nasals	wi		<i>m-</i>		<i>n-</i>				<i>ñ-</i>		
	wm		<i>-m-</i>		<i>-n-</i>		<i>(-ṇ-)</i>		<i>-ñ-</i>		
	≡	{/m/, (/v/ ^{borr} , /Ø/ ^{borr})}		{/n/, /ṇ/ ^{alt} }				{/ñ/}			
Sibilants	wi			<i>s-</i>				<i>ś-</i>		Glottal:	<i>h-</i>
	wm			<i>-s-</i>		<i>-ṣ-</i>		<i>-ś-</i>			<i>-h-</i>
	≡			{/s/ ^h , /ṣ/ ^{alt} , (/Ø/, /h/ ^{alt})}		{/ṣ/ ^h , /ṣ/ ^{alt} , (/ṣ/ ^{alt})}		{/ś/ ^h , /ṣ/ ^{alt} , (/Ø/, /h/ ^{alt})}		{/h/ ^h , (/Ø/, /ṣ/, /ṣ/ ^{alt})}	
Semivowels	wi		<i>v-</i>		<i>l-</i>		<i>r-</i>		<i>y-</i>		
	wm		<i>-v-</i>		<i>-l-</i>		<i>-r-</i>		<i>-y-</i>		
	≡	{/v/ ^h }		{/l/ ^h }		{/r/ ^h }		{/y/ ^h , /Ø/ ^{alt} , (/g/, /j/ ^h)}			
wi = word initial – most common; wm = word medial; > merged with voiced; ≡ common representations; () infrequent or not represented.											

- The nasal phonemes are *m*, *n*, and *ɳ*, while *ɳ̥* has been recorded as a phone only.
- There are three voiceless sibilants, *ʃ*, *ʂ* and *s*, along with voiced sibilants *ʒ* and *ʐ*.
- The semi-vowels *y*, *r* (dental), *l* and *w* occur in both medial and initial positions.

The phonological system of Torwali is more complex and varies from that of Gāndhārī and other MIA languages in a number of ways, including the retention of medial voiceless as well as voiced stops, the loss of aspirated voiced stops, and the emergence of classes of affricate consonants and voiced sibilants. A common feature shared between Gāndhārī and Torwali, and which differs from the other MIA languages (§ 9.2.1), is the retention of three unvoiced sibilants. The Gāndhārī record indicates that changes towards weaker pronunciation of the stops as approximants were occurring at that time, and would most likely have developed further towards elision, similar to the other MIA languages (§ 9.2.1). Such changes do not appear to have progressed in Torwali, suggesting a much slower and more stable pattern of change, while at the same time the language developed a more complex phonology, including a class of affricates, similar to other Dardic languages of the northwestern region (Bashir 2003: 821-823). Common retention of the three sibilants indicates an historical areal influence between Gāndhārī, Torwali and the other Dardic languages, however (Hock 2015: 122-124). Torwali effectively has a more conservative and complex phonological system compared to that of Gāndhārī, which was recorded some eighteen hundred years earlier. These differences make it difficult to reconcile any direct historical connection or development between the two languages, which is not assisted by the gap in the diachronic record.

Such considerations do not, however, preclude the possibility of an early connection between the languages, whose respective developments appear to reflect the significant topographical differences between the language locations despite their close geographical proximity. In this context Gāndhārī is a language of the lower valley plains and larger communities of Gandhāra, who were connected via the northwestern trade routes to the wider plains and MIA languages of northern South Asia. On the other hand, Torwali is a language of the upper Swat valley, spoken in a more remote high valley location, isolated from the main changes occurring to the Gāndhārī dialects spoken by the more widely connected communities of Gandhāra. As discussed in § 2.3, there is a tendency for languages in isolation to complexify or remain stable in low or limited contact situations, while languages tend to become simplified in high contact situations.

Hence it might be expected that Gāndhārī would have in time developed a simpler grammar and phonology like the other MIA languages, while perhaps retaining some conservative characteristics due to its relative isolation at the northwestern border region of the ‘plains.’ On the other hand, due to its more isolated high valley location, with presumed connections to other Dardic languages through mountain pass routes to Kohistan, Chitral and beyond, ancient Torwali has retained or developed a more complex grammar and phonology. At the same time, Gāndhārī appears to have absorbed some common areal features through contact with the Dardic languages to its immediate north. When considered in this way, and given their close geographical proximity, it is not unreasonable to suggest that Torwali could have developed from an early Gāndhārī dialect, or that both languages shared a common ancestor language, not dissimilar to Aśokan or pre-Aśokan Gāndhārī.

Another interesting feature of the Dardic languages is rhotic metathesis involving *rC* clusters, commonly referred to as “Dardic metathesis”²³⁷. While this has been recorded in a number of Dardic languages such as Kashmiri, Kalasha and Phalūra, it does not appear to be represented in Torwali. At the same time metathesis is recorded in many of the Gāndhārī sources, and appears to be more common in the Aśokan inscriptions than the later Buddhist manuscripts (§10.6), supporting suggestion of early areal connections between Aśokan or pre-Aśokan Gāndhārī and other northwestern languages.

²³⁷ E.g., Salomon 2000: 91; Allon 2001: 97; Lenz 2003: 45.

8.3 Changes to OIA consonant clusters in Gāndhārī

8.3.1 Stop plus clusters

8.3.1.1 Stop + stop clusters

Examples: Stop + stop clusters including geminates only occur word medially in OIA. Gāndhārī reflexes of OIA stop + stop clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.1.1a. These reflexes are almost always written as single unweakened stops in the original Kharoṣṭhī documents, which for reasons discussed below are interpreted to represent geminate stops in Gāndhārī. While the authors of earlier publications such as Hultsch (1925) and Konow (1929) have noted the assumed gemination without recording it as such, it has been indicated by the writing of a second bracketed consonant in many of the more recent publications, as in *bd* > *d(d)* or *bdh* > *(d)dh* for example. For the reasons discussed in § 8.1.2, a convention of bracketing the assimilated consonant in all geminate reflexes has been adopted for this thesis, in order to reflect the assimilation patterns and directions of linguistic influence. Hence the above example of *bd* > *d(d)* is recorded as *bd* > *(d)d* since the assimilated consonant is *b*.

Table 8.3.1.1a: Gāndhārī reflexes of stop + stop clusters including geminates

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-kk-	-(k)k-	BLS1-1	<i>vṛkkāḥ</i> > <i>broka</i>	1 CE
		BLS4-1	*-kokkuṭi- > -kokuḍi-	
		RS-5	<i>kukkuṭyāḥ</i> > <i>kuku[d](*)ia</i>	2 CE
-kt-	-(t)t-	Aś-H	<i>abhiśikta-</i> > <i>abhisita-</i>	3 BCE
		InsK-K	<i>ukta-</i> > <i>uta-</i> (Kurram pp. xciv, 155)	2 CE
		BLS1-1	<i>bhaktam</i> > <i>bhatu</i>	1 CE
		BLS4-1	<i>vimukti-</i> > <i>vimuti-</i>	
		BLS9-1	<i>viṣaktaḥ</i> > <i>[viṣat]o</i>	
		BCS5-1	<i>virakta-</i> > <i>virata-</i> , <i>-bhukteṇa</i> > <i>-bhuteṇa</i>	2 CE
		RS-5	<i>ananuyukta-</i> > <i>aṇuyu[t]a-</i>	2CE
		RS-20	<i>ukte</i> > <i>vute</i>	
	-(t)th-	BLS4-1	*viṣakti- > <i>visathi-</i> (#)	1 CE
-ggh-	-(g)g-	RS-12	P <i>taggha-</i> (Skt. * <i>tad-gha-</i>) > <i>taga-</i> (#)	2 CE

-cc-	-(c)c-	BLS4-1	<i>anuccāvacā-</i> > <i>aṇucavaga-</i>	1 CE
		RS-12	P <i>kaccit</i> > <i>kaca</i>	
-cch-	-(c)ch-	BLS1-1	<i>acchamb[h]ī-</i> > <i>acham̐bi-</i> , <i>icchati</i> > <i>ichati</i>	1 CE
		BLS1-2	<i>gacchāmi</i> > <i>ghachami</i> , <i>gacchatā</i> > <i>ghachateṇa</i>	
		RS-5	<i>icchā-</i> > <i>iche-</i>	2 CE
		RS-20	<i>gacchāmi</i> > <i>gachami</i>	
-jj-	-(j)j-	RS-12	P <i>vajjīnām</i> > <i>vajiṇa</i>	2 CE
		RS-20	<i>yāvajjīvam</i> (<i>yāvat-jīvam</i>) > <i>yava-jīva</i>	
-ḍg-	-rg-	BLS9-1	<i>khadḍga-</i> > <i>kharga</i> (#)	1 CE
-tt-	-(t)t-	BLS1-1	<i>citta-</i> > <i>cita-</i>	1 CE
		BLS9-1	(<i>upātivṛttaḥ</i>) > <i>uvativuto</i>	
		BLS4-1	<i>uttamaḥ</i> > <i>utama</i>	
		BCS5-2	<i>citta-</i> > <i>cita-</i> , <i>sampatti-</i> > <i>sapati-</i>	2 CE
		RS-5	<i>citta-</i> > <i>jita-</i>	2 CE
		RS-20	<i>āttamanāḥ</i> > <i>atamaṇo</i>	
	-tr-	BLS2-2	<i>uttīrṇaḥ</i> > <i>utrino</i> (#)	1 CE
		BCS5-2	<i>vittasukha-</i> > <i>vitrasua-</i>	2 CE
	-tv-	RS-12	<i>uttarim</i> > <i>utvara</i> (#)	2 CE
-tth-	-(t)th-	RS-12	<i>utthāya-</i> > <i>uṭṭhahi-</i> (#)	2 CE
-tk-	-(k)k-	BLS1-1	<i>satkr̥tya-</i> > <i>[sa]kacu-</i>	1 CE
		BLS1-2	* <i>utkubjaya-</i> > <i>ukuje</i>	
		BLS4-1	* <i>satkāya-</i> > <i>saka[y]a-</i>	
		RS-5	<i>utkaṇṭhati</i> > <i>ukaṭaḍi</i>	2 CE
-tp-	-(p)p-	Aś-H	<i>prajā-utpadane</i> > <i>pajupadane</i>	3 BCE
		BLS2-2	<i>satpuruṣa-</i> > <i>sapuruṣa-</i>	1 CE
		BLS4-1	* <i>utpanna-</i> > <i>upaṇa-</i>	
		InsK-K	<i>saṃutpāda-</i> > <i>saṃmupate-</i> (Kurram p. 155)	2 CE
		BCS5-1	<i>utpanna-</i> > <i>upaṇa-</i>	2 CE
		RS-5	<i>utpadyeta</i> > <i>[u]pajea</i>	
-ddh-	-(d)dh-	BLS1-1	<i>buddha-</i> > <i>budha-</i>	1 CE
		InsK-K	<i>buddha-</i> > <i>budha-</i> (Māṇikiāla pp. xciv, 149)	2 CE
		BCS5-1	<i>middha-</i> > <i>midha-</i>	2 CE
		RS-5	<i>śraddha-</i> > <i>ṣadha-</i>	
-dg-	-(g)g-	BLS4-1	BHS * <i>maudgalyāyana-</i> > <i>-[moga]laṇa-</i>	1 CE

		RS-5	<i>pudgala-</i> > [pu]gala-	2 CE
		RS-14	<i>pudgala-</i> > <i>pugala-</i>	
	-(g)gh-	BLS1-1	<i>-pudgala-</i> > <i>-pughala-</i> , <i>udga-</i> > <i>ugha-</i> (#)	1 CE
-dgh-	-(g)g-	BLS4-1	<i>udghāṭana-</i> > <i>ugaḍaṇa-</i> (#)	1 CE
-pt-	-(t)t-	Aś-H	<i>guptiḥ</i> > <i>guti</i> , <i>upta-</i> > <i>vuta-</i> , <i>naptāraḥ</i> > <i>natare</i>	3BCE
		BLS1-1	<i>anavapte</i> > <i>aṇodate</i> , <i>sapta-</i> > <i>sata-</i>	1 CE
		BLS2-2	<i>anuprāptaḥ</i> > <i>anuprato</i>	
		BLS4-1	<i>gupta-</i> > <i>guta-</i>	
		BLS9-1	<i>avakṣipta-</i> > [o]kṣita-	
		BCS5-1	<i>sapta-</i> > <i>sata-</i> , <i>vijñapti</i> > <i>viñati</i>	2 CE
		RS-5	<i>saptānām</i> > <i>satana</i>	
		RS-19	<i>saṃkṣipta-</i> > <i>sakṣita-</i>	
		RS-20	<i>ādīptāḥ</i> > <i>aḍita</i>	
	-tv-	BLS4-1	<i>sunidhyapta-</i> > <i>sunī[ja]tva-</i> (#)	1 CE
-bd-	-(d)d-	BLS1-1	<i>śabdam</i> > <i>chadu</i>	1 CE
	-(d)dh-	RS-12	<i>śabdam</i> > <i>śadha</i> (#)	2 CE
-bdh-	-(d)dh-	Aś-H	<i>labdha-</i> > <i>ladha-</i>	3 BCE
		BLS1-1	<i>labdha-</i> > <i>ladha-</i>	1 CE
		BLS1-2	BHS <i>praśrabdhi-</i> > <i>praśadha-</i>	
		BLS4-1	<i>labdha-</i> > <i>ladha-</i>	
		BLS9-1	<i>ārabdha-</i> > <i>aradha-</i>	
		BCS5-1	<i>pratīlabdha-</i> > <i>padīladha-</i>	2 CE
		RS-12	<i>-śrabdhyai</i> > <i>-ṣadhie</i>	
	-dhr-	BLS2-2	<i>pratīlabdham</i> > <i>padīladhro</i> (#)	1 CE
-bj-	-(j)j-	BLS1-2	<i>*utkubjaya-</i> > <i>ukuje-</i>	1 CE

Other OIA combinations not recorded in the Gāndhārī source documents include *kkh*, *kc*, *jjh*, *ṭṭ*, *ḍḍ*, *ḍḍh*, *ḍgh*, *dd*, *dg*, *db*, *dbh*, *pp*, *bb* and *bgh*²³⁸.

Overview: While a small number of atypical reflexes have been recorded, both unaspirated and aspirated OIA geminate stops are typically retained, and clusters of two stops undergo complete assimilation to a long version of the second element. In all cases the Gāndhārī reflexes are represented by the *akṣara* of a single stop in the Kharoṣṭhī script.

²³⁸ These are drawn from the tabulation of Sanskrit consonant clusters presented in Appendix A3. Some of these clusters, such as *-ḍḍ-* for example, are quite rare in occurrence.

Since the reflexes of OIA single intervocalic stops are generally recorded in modified forms which indicate lenition to fricatives or approximants (§ 8.2.1.2), the writing of unweakened single stops is interpreted as a representation of geminates, especially in situations where these might be expected. Further indications of geminates include the non-weakening of unvoiced phonemes such as the retention of *-(t)t-* instead of *-(d)d-*; the sporadic usage in some of the BL manuscripts of the diacritic which is normally used to mark pre-consonantal *r*, in locations which can only be interpreted as indicating a geminate, as for example in Skt. *khaḍga* > G *kharga* / P *khagga* (Salomon 1999: 122-123); and the presence of syllables ending in an assumed double consonant in the Dhp-G^K which are consistent with metrical lengths (Konow 1929: xcvi-xcviii)²³⁹. The writing of a single stop appears to be a consistent diachronic pattern from the third century BCE to the second century CE. Comparison with cognates is necessary to determine whether the original form was a geminate or stop-stop cluster.

Discussion: Dissimilar OIA stop + stop clusters form geminates through regressive place assimilation, which is common in the world's languages (§ 6.1). The majority of the OIA clusters possess a structural symmetry in the form of voiceless or voiced pairs such as *pt* / *tp*, *bd* / *db*, *kt* / *tk* and *dg* / *gdh*. The pronunciation of stop + stop clusters requires a mid-cluster adjustment between the two points of articulation, and the more distinct transition that occurs between the velar or labial and coronal stops may be a reason for these pairings. In all cases these assimilate in favour of the second consonant, including clusters such as *kt* > *tt* and *pt* > *tt*, where the initial velar and labial stops are less likely to be assimilated by the second coronal stop, suggesting that the vowel transition from the second stop is the most dominant perceptual cue in these cases (§§ 6.1.1, 6.1.2). In the recording of such geminates in Gāndhārī it is suggested that the first stop is bracketed to reflect these assimilation pattern, as in *kt* > *(t)t* for example, rather than *t(t)* as is sometimes the case in published materials (§ 8.1.2). Several types of atypical reflexes have also been recorded:

- Reflexes involving orthographic *r* are recorded in *ḍg* > *rg* in Skt. *khaḍga-* > G *kharga* (BLS9-1), *tt* > *tr* in *uttīrṇaḥ* > *utrino* (BLS2-2), *vittasukha-* > *vitrasua* (BCS5-2), and *bdh* > *dhr* in *pratilabdham* > *padiladhro* (BLS2-2). Pre- and post-consonantal *r* are typically indicated by diacritic conjuncts in the Kharoṣṭhī script (§§ 5.3.2.1, 5.3.2.2),

²³⁹ The form *-tt-* has been written with one *t* above the other by the Schoyen Scribe 2 (Glass 2000: 187), providing another small piece of evidence for the presence of geminates.

and in the absence of other indications of an *r* consonant, appears to have been used by BL scribe 2, BL scribe 9 and BC scribe 5 to indicate geminate consonants²⁴⁰.

- Atypical cases of aspiration are recorded in *kt* > (*t*)*th* in Skt. **viṣakti*- > G *visathi*- (BLS4-1), *dg* > (*g*)*gh* in *-pudgalam* > *-pughalu* and *udga*- > *ugha*- (BLS1-1), and *bd* > (*d*)*dh* in *śabdam* > *śadha* (RS-12). Baums (2009: 161) suggests that *visathi*- is an orthographic variant, while Silverlock (2015: 254-255) observes that *śadha*- is unusual in the context of a dialectic trend for the loss of distinction between aspirates and non-aspirates. The writing of *gh* for *g* is a common graphic habit of BL scribe 1 (§ 8.2.1.2.1).
- Deaspiration is recorded in *ggh* > (*g*)*g* in P *taggha*- > *taga* (RS-12), and *dgh* > (*g*)*g* in Skt. *udghāṭana*- > G *ugaḍaṇa*- (BLS4-1). These both involve an OIA cluster of the type -C*gh*-, and although they are the only recorded reflexes of such clusters, it is possible they are following the general trend for deaspiration of voiced aspirates (§ 8.2.1.2).
- Reflexes involving *tv* are recorded in *tt* > *tv* in Skt. *uttarim* > G *utvara* (RS-12), and *pt* > *tv* in *sunidhyapta*- > *sunī[ja]tva*- (BLS4-1). It is possible that the latter case could derive from a dialectical variation (Baums 2009: 161), while reflexes of *utvara*- / *utvari*- occur in several contexts in Gāndhārī, suggesting a borrowed or inherited form (Silverlock 2015: 253-4). Post consonantal *v* is typically recorded with a cursive upward flourish from the base to the right of the main akṣara (§ 5.3.2.1), and similar to use of the *r* diacritic could be seen as another way of indicating a geminate consonant.
- A reflex involving diacritic *ṭ* is recorded in the single example of *tth* > (*ṭ*)*ṭh* in Skt. *utthāya* > G *uṭhahi* (RS-12). The akṣara *ṭh* has a complex and possibly confused etymology²⁴¹. In this case suggested interpretations include either a dental or retroflex aspirate, *th* or *ṭh*, pronunciation as a de-aspirated retroflex geminate *ṭṭ* (Silverlock 2015: 254), while the diacritic could also be intended as a marker of gemination.

While the recording of Gāndhārī geminate stops using the graphemes for single stops can lead to the development of homographs in the written language, especially when conservative forms are used to represent the single stops, there appear to have been two

²⁴⁰ Salomon 1999: 122-123; 2000: 77, 88; Lenz 2010: 29. The case of *uttīrṇaḥ* > *utrino* (BLS2-2) may also involve metathesis of *r*.

²⁴¹ E.g., Brough 1962: 75-77; Baums 2009: 164-167; Silverlock 2015: 254. Retroflex articulation is common in MIA and reflects analogical extension from *tiṣṭhati* > *tiṭṭhati*

points of phonetic distinction between the single and geminate stops in the spoken language. Geminate stops are typically produced with a single but longer articulatory movement which contrasts in length with single stops (§ 4.6), while at the same time the pronunciation of intervocalic single stops was typically weakened in Gāndhārī (§ 8.2.1.2), and it is possible this phonetic contrast may have lessened the need to retain length contrasts in the diachronic development of the Gāndhārī geminates. While the interpretation of written single stops as geminates in Gāndhārī has been questioned based on comparison with modern Dardic languages (Grierson 1913: 141-144), this approach is not accepted by contemporary scholars for the reasons outlined above. Since the available Kharoṣṭhī source materials from Gandhāra only extend to around the fourth century CE, the diachronic record is limited in comparison to other MIA languages, and it is not possible to ascertain at what stage gemination was lost in descendant languages. This may however have occurred by around the tenth century CE, similar to other MIA languages (Misra 1967: 195).

Summary: a summary of the Gāndhārī reflexes of OIA stop + stop clusters is presented in Table 8.3.1.1b.

Table 8.3.1.1b: Summary of stop + stop cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-T ₁ T ₁ -/ /-T ₁ T ₂ -/	/-T ₁ T ₁ -/ /-T ₂ T ₂ -/ Suggested to record as <i>-(k)k-</i> etc.	[T ₁ :] [T ₂ :]	/T ₁ T ₁ / ≡ /T ₁ / ^{adv} /T ₂ T ₂ / ≡ {/T ₂ / ^{adv} , (/rT ₂) ^{alt} , (/T ₂ r/) ^{alt} , (/T ₂ ^h) ^{alt} } Geminate stops are typically represented by single unweakened stops. The <i>r</i> diacritics are occasionally used to represent gemination. Alternate aspirated forms appear to be orthographic variants.
/-T ₁ T ₂ ^h -/	/-T ₂ T ₂ ^h -/	[T ₂ T ₂ ^h]	/T ₂ T ₂ ^h / ≡ {/T ₂ ^h / ^{adv} , (/g/)} Unaspirated /g/ is an atypical reflex to OIA T ₁ .gh clusters.

8.3.1.2 Stop + nasal clusters

Examples: Stop + nasal clusters occur both initially and medially in OIA. Gāndhārī reflexes of OIA stop + nasal clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.1.2a.

Table 8.3.1.2a: Gāndhārī reflexes of stop + nasal clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-gn-	-g(g)-	BLS4-1	<i>agni-</i> > <i>agi-</i>	1 CE
		BCS5-1	<i>nagnaka-</i> > <i>ṇagao-</i>	
-jñ-	-ñ-	Aś-H	<i>jñati-</i> > <i>ñati-</i>	3 BCE
		InsK-M	<i>jñati-</i> > <i>ñati-</i> <i>jñati-</i> > <i>ñati-</i>	1 BCE 1 CE
		BLS1-1	<i>jñātinām</i> > <i>ñadi(*ṇa)</i> (p. 430)	1 CE
		BLS4-1	<i>jñātvā-</i> > <i>ñatva-</i> , <i>jñāna-</i> > <i>ñaṇa-</i>	
		BCS5-1	<i>jñāna-</i> > <i>ñaṇa-</i>	
	-(ñ)ñ-	Aś-H	<i>kr̥tajñatā-</i> > <i>kitrañata-</i> , <i>rājñā</i> > <i>raña</i>	3 BCE
		InsK-M	<i>ājñā-</i> > <i>-aṇa-</i>	1 BCE
		BLS1-1	<i>saṃjñām</i> > <i>saña</i>	1 CE
		BLS2-2	<i>abhijñāya</i> > <i>[a]bhiñae</i>	
		BLS2-1	<i>ājñata-</i> > <i>añada-</i>	
		BLS4-1	<i>prajñā-</i> > <i>praña-</i> , <i>vijñāna-</i> > <i>viñaṇa-</i>	
		BLS9-1	<i>kr̥tajñatā-</i> > <i>kitañata-</i>	
		InsK-M	<i>vijñāna-</i> > <i>viñana-</i>	2 CE
		BCS5-1	<i>aprajñapti-</i> > <i>aprañati-</i>	
		RS-5	<i>saṃjñā-</i> > <i>ṣaṇa-</i>	
		RS-19	<i>abhijñā-</i> > <i>aviñae</i>	
		RS-20	<i>prajñayā</i> > <i>pañae</i>	
-tm-	-tm- -tv-	InsK-M	<i>ātman-</i> > <i>atmaṇa-</i> (#)	2 CE
		Aś-H	<i>ātman-</i> > <i>atva-</i> (Mānsehra)	3 BCE
		InsK-M	<i>ātman-</i> > <i>atvaṇa-</i>	1 CE
		BLS1-1	<i>ātmakāmena</i> > <i>[atva]ghamena</i>	
		BLS2-1	<i>ātmanā</i> > <i>[atva]n[o]</i>	
		BLS2-2	<i>anātmān-</i> > <i>anatva-</i>	

		BLS4-1	<i>adhyātman-</i> > <i>ajatva-</i> , <i>ātman-</i> > <i>atva-</i>	
		BCS5-1	<i>ātman-</i> > <i>atva-</i>	
		RS-5	<i>ātmanyam</i> > <i>atvaṇi[a]</i>	
		RS-19	<i>prahitātmā-</i> > <i>pahidatva-</i>	
		RS-20	<i>adhyātman</i> > <i>acatva</i>	
	-t(t)-	Aś-H	<i>ātman-</i> > <i>ata-</i> (#)	3 BCE
		InsK-M	<i>ātman-</i> > <i>ata-</i> (#)	2 CE
	-(p)p-	BLS2-1	<i>ātman-</i> > <i>apanaga-</i> (#)	1 CE
		InsK-M	<i>ātmanā</i> > <i>apano</i> (#)	2 CE
-dm-	-dum-	BLS4-1	<i>padmaḥ</i> > <i>paduma</i>	1 CE
-dhm-	-dhum-	BLS1-1	<i>uddhmātakam</i> > <i>udh[u]maï</i>	1 CE
-pn-	-pun-	Aś-H	<i>prāpnoti</i> > <i>prapuṇati</i>	3 BCE

Other OIA combinations not recorded in the Gāndhārī source documents are *kn*, *kṇ*, *km*, *khn*, *gm*, *ghn*, *ghm*, *cñ*, *cm*, *chn*, *chm*, *jm*, *tn*, *thn*, *dn*, *dhn*, *ṭṇ*, *ṭm*, *ḍn*, *pm*, *bhn*, *bhṇ* and *bhm*.

Overview: Gāndhārī reflexes are only recorded for six of the thirty OIA stop-nasal combinations, and the diachronic record is limited except in the case of *jñ* > (*ñ*)*ñ* and *tm* > *tv* which are quite widely attested. Other changes include the apparent gemination of *gn* > *g(g)* and in one instance of *tm* > *t(t)*, while *dm* > *dum*, *dhm* > *dhum* and *pn* > *pun* all involve insertion of an epenthetic *u* vowel. Where applicable gemination is assumed and indicated in the table by the addition of a bracketed consonant in place of the assimilated consonant, for the reasons discussed in § 8.1.2. While other reflexes are recorded in the Kharoṣṭhī script using single *akṣaras*, the *tv* cluster is written by the addition of a diacritic flourish representing *v* following the main *akṣara* (§ 5.3.2.1).

Discussion: While stop + nasal clusters have increasing sonority, difficulties of articulatory co-ordination can inhibit their use as syllable onsets, and among the small number of these clusters that occur word initially in OIA, only the *jñ*- cluster is used on a regular basis (Appendix A4). Reflexes of the OIA stop-nasal clusters recorded in the Gāndhārī source materials appear to develop in four different ways, as described below:

- Development of the homorganic palatal clusters *jñ*- > *ñ*- and *-jñ*- > *-(ñ)ñ*- is well attested diachronically in Gāndhārī. The medial geminate is also recorded in reflexes of OIA -Ny- clusters, as in *-{n. ṇ}y-* > *-(ñ)ñ-* or *-ṇ.ñ-* (§ 8.3.2.4.3). While the medial

reflex is represented as the geminate nasal *-ñ(ñ)-* in many of the source publications, it is typically recorded as a single *akṣara* in the Kharoṣṭhī script. The geminate is interpreted based on parallels in other MIA languages, and the inclusion of the *anusvāra* diacritic in some of the *-Ny-* reflexes. While a conjunct *akṣara* representing the *jñ* cluster has been identified in one of the third century CE Bamiyan palm leaf fragments, indicating an awareness of the OIA cluster, it has been suggested that this is an incorrect Sanskritisation of the Gāndhārī manuscript in what is a single case (Glass 2000: 129). The word medial cases develop through the spread of nasality, which occurs through regressive nasalisation in the palatal cluster as *jñ > ññ*, and progressive nasalisation following anticipatory palatalisation of the coronal nasal in *{ny, ñy} > ñy > ññ*. In both cases the palatal nasal *ñ* appears to control and cause nasalisation of the conjunct palatals *j(ñ)* and *(ñ)y*. The pronunciation of word and syllable initial *jñ-* in Sanskrit presents articulatory difficulties, and the uncertain pronunciation of this cluster appears to have been resolved in different ways in the modern IA languages (Appendix A4). In word initial positions it is possible this was pronounced as a unitary ‘pre-stopped’ nasal segment, where the orthographic *j-* may simply serve as a place signal for the nasal. In this approach there may be little phonetic distinction between Sanskrit *jñ-* and Gāndhārī *ñ-*, which may also apply if Sanskrit medial *-jñ-* was syllabified as an onset. In medial positions the V₁C₁ formant transitions may provide stronger perceptual place cues for the stop, which may be rendered in the coda of the preceding syllable in OIA.

The inclusion of word initial *ñ-* and medial *-ññ-* suggests the development of *ñ* as an unconditioned phoneme, and it is necessary to address the phonemic status of these reflexes in Gāndhārī. While the three nasal geminates *-nn-*, *-ññ-* and *-ññ-* were represented orthographically at different times in the central and eastern branches of MIA, it is suggested these had coalesced into a single sound as the dental nasal geminate *-nn-* (Misra 1967: 139-141). At the same time, two contrastive geminate nasals appear to have been retained in the other branches, represented by *-nn-/ññ-* and *-ññ-*. In Gāndhārī the use of graphic *ñ* appears to have been retained consistently until at least into the second century CE, and any graphic merger of palatal *-ññ-* with dental *-nn-* is not evident in the source materials. Some near minimal pairs which contrast on the basis of *ñ-* / *-ññ-* and *n/ñ* are listed below, suggesting that the *ññ* clusters recorded in Gāndhārī were at least to some extent phonemic.

Sanskrit <i>jñ</i> > Gāndhārī <i>(ñ)ñ</i>	Sanskrit <i>n</i> > Gāndhārī <i>n/ṇ</i>
<i>jñātiñ</i> > <i>ñati-</i> or <i>ñadi-</i> "kinsmen" (Salomon 2000: 229; 2008: 430)	<i>nadī-</i> > <i>nadi-</i> "river" (Lenz 2010: 173) <i>nandi</i> > <i>ṇadi-</i> "relish" (Baums 2009: 649).
Sanskrit <i>ny</i> > Gāndhārī <i>ñ(ñ)</i>	Sanskrit <i>n</i> > Gāndhārī <i>n/ṇ</i>
<i>puṇya-</i> > <i>puñ(ñ)a-</i> "merit" (Baums 2009: 656)	<i>punaḥpunaḥ</i> > <i>puṇapuṇo</i> "again and again" (e.g., Baums 2009: 656)
<i>anya-</i> > <i>añ(ñ)a-</i> "other" (Baums 2009: 619)	<i>anna-</i> > <i>an(n)a-</i> "food" (Baums 2009: 621)
<i>śūnya-</i> > <i>śuñ(ñ)a-</i> "empty" (Glass 2007: 248)	<i>śuna-</i> > <i>śuna-</i> "dog" (GD s.v.)

Since palatal *ñ* mostly arises as a conditioned allophone in homorganic clusters in Sanskrit (§ 4.3), the apparent emergence of palatal *ñ*- or *-ññ* - as independent nasal phonemes in Gāndhārī could be interpreted as the development of a new phoneme. On the other hand, homonym pairs would develop in the above examples if both the palatal nasality and gemination of the *-ññ*- reflexes were lost. Since nasalisation of the *j* and *y* palatals had altered the nature of the original conditioning element, it is possible that a phonetically simpler expression of the nasal reflex may have emerged in time, including shortening of any gemination. While the above cases only represent a small sample, the number of similar cases appears to be limited, and the meaning contrasts can mostly be interpreted contextually, or, as discussed in § 8.2.1.2.6.2, be adjusted for in other ways in the diachronic development of the language. In this context it is expected that the value of maintaining such limited phonemic contrasts in the spoken language would have diminished in time, especially if these were based on the development of new contrastive phoneme with limited application. This suggests that palatal *ñ* was largely if not totally non-phonemic in Gāndhārī, similar to eastern and central MIA, and continued use of the *ñ akṣara* was possibly a scholarly orthographic conservatism. Although difficult to establish, it is also possible that orthographic *ñ*- was used as a shorthand representation for the word initial *jñ*- cluster.

- The development of medial *-tm-* > *-tv-* is regularly attested in the manuscript and inscriptional sources in the common example of Skt. *ātman-* > G *atva-*. This involves loss of nasalisation along with deocclusion in anticipation of the following vowel, and a staged development of *tm* > *tp* > *tv* might be expected. Although the intermediate stage is not recorded, the assimilation of *t* by the *p*, as in *tm* > *tp* > *pp*, is recorded in

two examples, which could however also be borrowings. Oberlies (2019: 152) has suggested that **tv* may be an intermediate stage in the full assimilation of the labial nasal by the coronal stop in Skt. *ātman-* > P *attan-*. The reflex of *tm* > *t(t)* also occurs in some of the other Aśokan inscriptions (Mehendale 1948: 26), suggesting that its use in the Shāhbāzgarhī inscription (Hultzschi 1925: lxxxviii) may be an imported form. The single *tm* reflex is assumed to be a conservative form or palaeographic error.

- The development of medial Skt. *-gn-* > G *-gg-* is only recorded in two examples, and suggests progressive place and manner assimilation of the coronal nasal to the preceding velar stop. While contrary to the normal tendencies for regressive place assimilation to occur, velar stops possess stronger perceptual place cues than coronal nasals (§ 6.1.1). This reflex is also similar to that which occurs in other MIA languages (Table C2 [Appendix C]), suggesting the possibility of borrowed forms.
- In limited examples, the medial *-dm-* and *-pn-* clusters appear to develop through epenthetic insertion of the labial vowel *-u-*, although this may represent allophonic pronunciation of these clusters with labial consonants. The development of *dm* > *dum* only appears in *paduma-* as a loanword from another MIA language (Baums 2009: 180), while descendant languages show a development of *dm* > *mm* which is consistent with tendencies for regressive assimilation of the coronal stop by the labial nasal. The development of the present tense third person singular form of the OIA verbal root \sqrt{ap} is recorded in Skt. *pra + āpnoti* > G *pra-punati* in the Aśokan inscriptions.

Summary: Gāndhārī reflexes of only two of the OIA stop + nasal clusters are attested on a consistent basis, and a summary of these is presented in Table 8.3.1.2b.

Table 8.3.1.2b: Summary of stop + nasal cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
<i>/jñ-/</i> <i>/-jñ-/</i>	<i>/ñ-/</i> <i>/-ññ-/</i>	[ɲ-] [-ɲ:-]	<i>/ñ-/</i> ≡ <i>/ñ-/</i> ^{adv} <i>/-ññ-/</i> ≡ <i>/-ñ-/</i> ^{adv} Gāndhārī geminate palatal nasals are represented by the single nasal character <i>/ñ/</i> .
<i>/-tm-/</i>	<i>/-tv-/</i>	[-tv-]	<i>/-tv-/</i> ≡ <i>/-tv-/</i> ^{adv}

8.3.1.3 Stop + fricative clusters

Examples: Stop + fricative clusters occur both initially and medially in OIA. Gāndhārī reflexes of OIA stop + fricative clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.1.3a.

Table 8.3.1.3a: Gāndhārī reflexes of stop + fricative clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī		Era
		For source references refer to Table 8.1		
		* = unattested equivalent (#) atypical reflexes		
-kś-	-kṣ-	BLS4-1	*ākśāti > aikṣadi (Baums 2009: 618)	1 CE
kṣ-	kṣ-	Aś-H	kṣaṇati > kṣaṇati	3 BCE
		InsK-M	kṣatrapa- > kṣatrapa-	1 BCE
		BLS1-1	kṣapitāḥ > kṣavida	1 CE
		BLS9-1	kṣudhām > kṣuṣa	
		BLS2-2	kṣiṇāti > kṣinedi	
		InsK-M	kṣaṇa- > kṣana-	2 CE
		RS-5	kṣayam > kṣao	2 CE
	kh-	RS-12	kṣamaṇīyam > khamañio	
		InsK-M	kṣaṇa- > khaṇa-	2 CE
-kṣ-	-kṣ-	Aś-H	akṣati > akṣati	3 BCE
		BLS1-1	tat-kṣaṇe > takṣaṇo, dakṣiṇā > dhakṣiṇa	1 CE
		BLS2-2	lakṣam > lakṣo	
		BLS4-1	cakṣus- > cakṣu-, pakṣa- > pakṣa-	
		BLS9-1	avakṣipta- > [o]kṣita-	
		RS-5	vṛkṣa- > rukṣa-	2 CE
		RS-12	nikṣipyā- > ṇikṣivita-	
		RS-14	-yakṣa- > -yakṣa-	
		InsK-M	Takṣaśilā- > Takhaśila-	1 BCE
			bhikṣu- > bhikhu-	1 CE
		BLS1-1	bhikṣu- > bhikhu-	1 CE
		BLS1-3	bhikṣu- > bhikhu-	
		BLS4-1	aśaikṣatā > aśekhada, cakṣu- > cakhu-	
		BLS9-1	apekṣām > uvekha, -cakṣuḥ > -cakhu	

		RS-5	<i>bhikṣuḥ</i> > <i>bhikkhu</i>	2 CE
	<i>[-(c)ch-]</i>	BLS1-1	<i>kukṣau</i> > <i>ku(c)chie</i> (#)	
<i>-ts-</i>	<i>-ts-</i>	BLS1-1	<i>*jighatsābhiḥ</i> > <i>jihitsahi</i>	
		BLS9-1	<i>vicikitsām</i> > <i>(*vi)yigitsa</i>	
		BLS4-1	<i>amatsarī-</i> > <i>amatsari-</i>	
	<i>-tś-</i>	InsK-M	<i>saṃvatsara-</i> > <i>saṃvatśara-</i> (#) palatalisation	1 BCE -2 CE
	<i>-s-</i>	InsK-M	<i>utsava-</i> > <i>usava-</i> (#)	
<i>-ps-</i>	<i>-ś-</i>	BCS5-1	<i>*jugupsitavya-</i> > <i>[j]uhosidave</i>	2 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *ks*, *ts*, *tś*, *tś* and *pś*.

Overview: The majority of recorded reflexes are for the OIA *kṣ* and *ts* clusters, which are generally retained. The reflex of *kṣ* > *kh* also occurs on a regular basis, most commonly in *bhikṣu-* > *bhikhu-*, which is most likely a retained eastern form due to the status of the word. Since this reflex is typically written as an unweakened intervocalic stop, it is assumed to be a geminate (§ 8.3.1.1). The Kharoṣṭhī script includes a unique ligature to record the *kṣ* cluster (𑀓𑀲) (§ 5.3.3), while *ts* (𑀓𑀲𑀭) is recorded as a conjunct formed between the two component akṣaras (§ 5.3.4).

Discussion: The retention of the phonological *kṣ-* cluster suggests that the fricative release of the stop is sufficiently stable for this cluster to retain its phonological status as a TS type cluster, although similar to OIA (Appendix A3) the actual pronunciation in Gāndhārī remains unclear²⁴². Based on analysis of loanwords and other evidence a phonetic development of *[kṣ]* > *[tṣ]* has been suggested, involving regressive place assimilation of the velar stop to the coronal retroflex position²⁴³, as in *[yakṣa-* > *yaṭṣa-]* for example. This is consistent with the relative stability of fricatives compared with plosives (Appendix B), and the observed stability of the homorganic dental-alveolar cluster *ts* [ts], as in Skt. *amatsarī-* > G *amatsari-* for example. It has been suggested that OIA *kṣ* and *ts* may have acquired the status of single phoneme affricates (Kobayashi 2004: 59-60), while affricate pronunciation of the Gāndhārī palatal stops *c* and *ch* has also been suggested (§ 8.2.1.2.2). In this context the recording of Skt. *kukṣau* > G *kuchie* offers a further hint of affrication.

²⁴² Brough 1962: 72-73; Salomon 2008: 123.

²⁴³ Brough 1962: 72-73; Baums 2009: 168-169.

The recording of the velar aspirate in $k\varsigma > (k)kh$, which reflects a process of debuccalisation, is largely seen in technical terms which are likely to be borrowings²⁴⁴.

Summary: a summary of the Gāndhārī reflexes of OIA stop + fricative clusters is presented in Table 8.3.1.3b.

Table 8.3.1.3b: Summary of stop + fricative cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
$/k\varsigma-/$ $/-k\varsigma-/$	$/k\varsigma-/$ $/-k\varsigma-/$	$\{[k\varsigma],$ $[k\widehat{\varsigma}];$ $[t\varsigma],$ $[t\widehat{\varsigma}]\}$	$/k\varsigma-/ \equiv \{ /k\varsigma-/^h, (/kh-/)^{alt} \}$ $/-k\varsigma-/ \equiv \{ /-k\varsigma-/^h, (/kh-/)^{alt} \}$ These forms are stable. The alternate forms are found in technical terms attributable to a source dialect.
$/-ts-/$	$/-ts-/$	$\{[ts],$ $[t\widehat{s}]\}$	$/-ts-/ \equiv \{ /-ts-/^h, (/tś-/^{alt}, /-s-/^{alt}) \}$ These forms are stable. The alternate forms appear to be orthographic variants.

8.3.1.4 Stop + semi-vowel clusters

The OIA semi-vowels are traditionally listed in the order of place of articulation, as in the palatal y , alveolar r , dental l and labio-dental v . Apart from the rhotic r , which is classified as an alveolar tap, the other semi-vowels are approximants, with l classified more specifically as a lateral approximant (§ 4.1). Apart from stop + l , clusters of stop + semi-vowel are common in OIA (Appendices A3, A4). They have rising sonority and form syllable onsets in many initial as well as medial positions. The following sections follow the assumed strength hierarchy of the semi-vowels (§ 6.1.3), rather than the traditional place order.

8.3.1.4.1 Stop + l clusters

Examples: Stop + l clusters have a limited distribution in OIA. Only a small number of reflexes have been recorded in Gāndhārī, and these are presented in Table 8.3.1.4.1a.

²⁴⁴ Brough 1962: 102; Salomon 2008: 124; Baums 2009: 169.

Table 8.3.1.4.1a: Gāndhārī reflexes of stop + *l* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>kl-</i>	<i>kil-</i>	RS-12	<i>klāmatha-</i> > <i>kilamaśa-</i>	2 CE
	<i>kr-</i>	RS-5	<i>klomakam</i> > <i>[kr]umao</i> complex etymology	
<i>-kl-</i>	<i>-kil-</i>	BLS9-1	<i>upakleśān</i> > <i>uvakileśa</i>	1 CE
	<i>-kr-</i>	BLS4-1	<i>śukla-</i> / <i>śukra-</i> > <i>śukra-</i> (Baums 2009: 676)	
<i>gl-</i>	<i>gil-</i>	BLS2-1	<i>glānakaḥ</i> > <i>gilanago diss w- ep v</i>	1 CE
	<i>ghil-</i>	BLS1-1	<i>glānaḥ</i> > <i>[ghil](*)aṇo</i>	
<i>pl-</i>	<i>pr-</i>	RS-5	<i>plīhakaḥ</i> > <i>priao</i> complex etymology	2 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *chl*, *bl* and *bhl*.

Overview: Based on the limited sample these clusters appear to remain stable. Reflexes include replacement of the approximant *l* as the tap *r*, or dissolution of the cluster by insertion of an epenthetic vowel. These reflexes are recorded both initially and medially. There does not appear to be any unique ligature, combined character or diacritic used to represent stop + *l* clusters in the Kharoṣṭhī script (Glass 2000: 114-126, 160).

Discussion: Anaptyctic insertion is not uncommon in Gāndhārī, and appears to represent the recording of allophonic developments within clusters in most instances (§ 10.1.1). The vowel qualities are typically determined by features of the second consonant, and in this case the epenthetic vowel *-i-* develops in anticipation of the following alveolar lateral *l*, and the phonetic insertion of $\emptyset > i / k_l$ is suggested as a narrower allophonic transcription of the *kl* cluster. This may or may not represent a developing morphological change. The alternate reflexes of *kl* > *kr* and *pl* > *pr* appear to represent a change from the alveolar lateral to a tap in the same place of articulation. The examples of Skt. *klomakam* > G *[kr]umao* and *plīhakaḥ* > *priao* occurring in the RS-5 manuscript are technical terms related to body parts which have a complex etymology (Glass 2007: 122, 157-158), although these changes could also reflect influence from a northwestern "r-dialect." Laterals and rhotics share some common features and are often classed together as 'liquids', while in some languages a single underlying phoneme can be pronounced as either a lateral or rhotic (Ladefoged and Maddieson 1996: 182-183, 243), which may be

the reason that alternate spellings such as *śukla-* / *śukra-* have emerged in Sanskrit. Based on the small sample, the *Cr* reflexes, including *śukla-* / *śukra-* > *śukra-*, appear to be associated with the presence of a labial vowel or consonant. There is a wide articulatory variation among the family of rhotics (Ladefoged and Maddieson 1996: 244), and it is difficult to establish which conditioning features may cause such alternations. The approximant alveolar rhotic [ɹ], which shares both place and some manner features with the alveolar lateral approximant [l], has a relatively low third formant, and in some instances is pronounced with a narrowing of the lips²⁴⁵, features shared by *p* and the back labial vowels *u* and *o*. This suggests a possible influence of these features onto *l* from the adjacent labial vowel or consonant, resulting in allophonic pronunciation as the alveolar approximant rhotic [ɹ], but recorded using the phoneme *r*.

The development of these clusters in Gāndhārī varies according to the phonetic environment, and this variability suggests that the recorded reflexes most likely represent allophones. In the absence of more examples it is unclear whether unchanged conservative forms were also recorded in the language, and the extent to which these allophones are transitional in a developing sound change. For these reasons, and in the absence of a wider data base, it is suggested that conservative forms are retained as the indicative reflexes.

Summary: a summary of the Gāndhārī reflexes of OIA stop + *l* clusters is presented in Table 8.3.1.4.1b. This tabulation assumes that the clusters have been recorded as transitional allophones, rather than being a fully developed sound change.

Table 8.3.1.4.1b: Summary of stop + *l* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/ Tl-/	/ Tl-/	[Tɹ] / labial	/Tl/ ≡ {/Tr/ ^{tr} , /Til/ ^{tr} }
/-Tl-/	/-Tl-/	[Til] / otherwise	Gāndhārī /Tl/ clusters are typically represented by the allophones /Tr/ ^{tr} and /Til/ ^{tr} , which arise in labial and non-labial environments respectively.

²⁴⁵ Kent and Read 1992: 138-141; Ladefoged and Maddieson 1996: 234-235.

8.3.1.4.2 Stop + v clusters

Examples: Gāndhārī reflexes of OIA stop + v clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.1.4.2a.

Table 8.3.1.4.2a: Gāndhārī reflexes of stop + v clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>kv-</i>	<i>kv-</i>	BLS1-1	* <i>kvathya-</i> > <i>kvachia-</i>	1 CE
	<i>kuv-</i>	BLS2-2	<i>kvacit</i> > <i>kuvaci</i>	
<i>-jv-</i>	<i>-c- -ĉ-</i>	RS-20	<i>samprajvalitaḥ</i> (pp) > <i>sapacalida</i> / <i>sapaĉalida</i>	2 CE
	<i>-j̄-</i>		<i>samprajvalitaḥ</i> > <i>sapaĵalida</i>	
<i>tv-</i> <i>-tv-</i>	<i>tv-</i>	BLS1-3	<i>tvacam</i> > <i>tvaya</i> (also <i>śpaya</i> – scribal error)	1 CE
		RS-5	<i>tvaca</i> > <i>tvaya</i>	2 CE
	<i>-tv-</i>	Aś-H	<i>tadātvam</i> > <i>tadatvaye</i> ,	1 BCE
		InsK-M	<i>sattva-</i> > <i>satva-</i>	1 CE
		BLS1-1	<i>saṃtvaritaḥ</i> > <i>satvaridu</i>	1 CE
		BLS1-2	<i>catvāra-</i> > <i>catvara-</i>	
		BLS4-1	<i>pūritatvāt</i> > <i>puridatva</i> and others	
		RS-5	<i>abhāvitatvāt</i> > <i>abhavidatva</i>	2 CE
		RS-20	<i>adṛṣṭatvāt</i> > <i>adhriṭhatva</i>	
		InsK-M	<i>sattva-</i> > <i>satva-</i>	2 CE
		InsK-M	<i>sattva-</i> > <i>satva-</i>	3 CE
		BLS1-2	<i>yakṣatvam</i> > <i>yakṣatu</i> abstract suffix <i>-tva</i>	1 CE
		InsK-M	<i>sattva-</i> > <i>sapa-</i>	1 CE
		InsK-M	<i>ekacatvārimśa-</i> > <i>ekacapariśa-</i>	2 CE
		BLS1-1	<i>arhattvam</i> > <i>arahapa</i> (<i>ttv</i> > <i>p</i>)	1 CE
		BLS4-1	<i>arhattva-</i> > <i>[ra]hapa-</i>	
<i>-tv-</i> Absolutives <i>-tvā</i> Generally <i>-itvā</i>	<i>-tv-</i>	BLS1-1	<i>gatvā</i> > <i>ghatva</i> and others	1 CE
		BLS1-2	<i>gatvā</i> > <i>ghatva</i> , <i>śrutvā</i> > <i>ṣutvaṇa</i>	
		BLS9-1	<i>-tvā</i> > <i>-tva</i>	
	<i>-t(t)-</i>	RS-14	<i>gatvā</i> > <i>gha[tva]</i>	2 CE
		Aś-H	<i>-tvi</i> > <i>-ti</i> typical	3 BCE
		InsK-M	<i>kṛtvā</i> > <i>karita</i>	1 CE
		BLS1-1	<i>ujjhītvā</i> > <i>ujita</i> , * <i>karītvā</i> > <i>karita</i> and others	1 CE

		BLS1-2	<i>-itvā</i> > <i>-ita</i> several examples	
		BLS1-3	<i>dālayitvā</i> > (<i>*da</i>) <i>laiṭa</i>	
		BLS4-1	<i>bāhitvā</i> > <i>vahita</i> and others	
		BLS9-1	<i>-itvā</i> > <i>-ita</i> several	
		RS-5	<i>dāmtvā</i> > <i>damita</i> several	2 CE
		RS-20	<i>-itvā</i> > <i>-it(*a)</i>	
<i>dv-</i>	<i>dv-</i>	RS-12	<i>dvādaśa-</i> > <i>dvaḍaśa-</i>	2 CE
		RS-20	<i>dvāraiḥ</i> > <i>dvarehi</i> , (<i>śaḍ-dvāram</i> > <i>sa-dvaro</i>)	
	<i>duv-</i>	Aś-H	<i>dvau</i> > <i>duv[i]</i>	3 BCE
		Aś-M	<i>dvi-</i> > <i>duvi-</i> , <i>duve</i> (Mānsehra)	
	<i>du-</i>	BLS2-2	<i>dve</i> > <i>[du]ve</i>	1 CE
		BLS2-2	<i>dviprakārā-</i> > <i>dupragara-</i>	1 CE
		BCS5-1	<i>dve</i> > <i>[du]e</i>	
		RS-12	<i>dvītyam</i> > <i>d(*u)ḍie</i>	2 CE
	<i>d-</i>	Aś-H	<i>*dvi-ardha-</i> > <i>diadha-</i>	3 BCE
	<i>b-</i>	Aś-H	<i>dvādaśa-</i> > <i>badaya[śa?]-</i>	3 BCE
		BLS2-2	<i>dvitīyaḥ</i> > <i>bidiga</i> , <i>dvitīyam</i> > <i>bidige</i>	1 CE
		RS-5	<i>dvādaśa-</i> > <i>[bada]śa-</i>	2 CE
		RS-12	<i>dvitīyam</i> > <i>biḍi</i>	
	<i>rv-</i>	BLS1-1	<i>dvāram</i> > <i>rvara</i> , also <i>bharadvājah</i> > <i>bharvayo</i>	1 CE
<i>-ddhv-</i>	<i>-dhv-</i>	BLS4-1	<i>*vrddhvā</i> > <i>vudhva</i>	1 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *gv*, *ghv*, *chv*, *ṭv*, *ḍv*, *ḍhv*, *thv*, *pv*, *bv* and *bhv*, although some of these only occur infrequently.

Overview: The OIA stop + *v* clusters are generally retained in Gāndhārī, although alternate reflexes are recorded in some cases. These include single stop reflexes in the case of *-jv-* clusters and Sanskrit *-itvā* absolute endings, and the dissolution of some *dv-* clusters with an epenthetic *-u-* vowel. Written single stop reflexes are assumed to mark geminates (§ 8.3.1.1). In the Kharoṣṭhī script a stop followed by the approximant *v* is typically recorded with a cursive flourish extended upward from the base to the right of the main *akṣara*, as in *ta* (𑀲𑀸𑀓) and *tva* (𑀲𑀸𑀓𑀲𑀸𑀓) for example (§ 5.3.2.1).

Discussion: The *tv* cluster is the most widely attested OIA stop + *v* cluster in Gāndhārī. It is recorded twice in word initial examples where it remains stable, and more widely in

medial positions where it is mostly stable, but single stop reflexes of *-t-* and *-p-* are also recorded. Like the other semi-vowels *v* has a well defined formant structure giving it a robust identity which is less susceptible to progressive assimilation by a preceding consonant (§ 6.1), although it is also prone to lenition in the articulatory transition from a stop to a following vowel. The development as a medial geminate stop, similar to the MIA reflex (§ 9.1.2), is mostly seen in Sanskrit absolutes involving the linking vowel *-i-*, as in Skt. *ujjihitvā* > G *ujit(t)a* and *bāhitvā* > *vahit(t)a* for example. This occurs consistently in the diachronic record commencing from the Aśokan inscriptions, and appears to be a phonological or graphic alternation peculiar to Gāndhārī²⁴⁶. A less frequently recorded reflex is the development of *-tv-* > *-p(p)-*, which is also attested in the Central Asian Niya documents²⁴⁷. The labialisation may reflect a two-stage process of fortition of *v* > *p* followed by regressive place assimilation by the labial stop, although no intermediate stages are recorded. It has been suggested that the cases of Skt. *arhattva-* > G. *rahap(p)a-* may be triggered by the syllable boundary within the OIA *-t.tv-* cluster (Baums 2009: 180). The one-off recording of *-t-* in the Aśokan inscriptions appears to be a borrowed form, which is common to the inscriptions (Mehendale 1948: 231 [§ 412 (iv)]).

Many word initial reflexes of the OIA *dv* cluster are recorded in various forms of the Sanskrit numerals *dvi-* (“two”) and *dvādaśa-* (“twelve”), as well as in *dvāra-* (“door”). Developments include two cases in which the cluster is retained, others in which the cluster is reduced by insertion of epenthetic *-u-*, and several involving labialisation to *b-*. Many of the epenthetic vowels recorded in Gandhārī appear to be allophones which represent phonetic developments occurring in particular environments (§ 10.1.1), such as *Ø* > *u* / *d_v* in this case, and the extent to which these represent a developing morphological change is uncertain. The intermixing of *duv-* and *d-* appears to be a common pattern in the Aśokan inscriptions (Mehendale 1948: 231, § 412 (v)), suggesting that *d-* is a borrowed form in Gāndhārī. The development of *dv-* > *b-* suggests labialisation similar to the medial *-tv-* > *-p-* case, and while its attestation in the Central Asian and MIA records points toward a borrowed form, this may also be an advanced form (Silverlock 2015: 261). The recording of *dv-* > *rv-* in Skt. *bharadvājaḥ* > G *bhar.vayo* suggests a shortening of *bharad.* > *bhar.* with the syllable boundary retained, with a possible misinterpretation of this reflex in *dvāram* > *rvara*. While only seen in limited word initial

²⁴⁶ Salomon 2000: 89; Baums 2009: 180.

²⁴⁷ Burrow 1937: 16; Salomon 2008: 122.

examples, the *kv-* cluster remains stable or develops as epenthetic *kuv-*, while *-ddhv-* > *-(d)dhv-* also remains stable in the one case. The recording of *-jv-* > {-*j̄-*, *-c-*, *-c̄-*} suggests palatalisation and regressive assimilation as a geminate palatal stop reflex.

Summary: a summary of the more widely attested Gāndhārī reflexes of OIA stop + *v* clusters is presented in Table 8.3.1.4.2b.

Table 8.3.1.4.2b: Summary of stop + *v* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
<i>/tv-/</i> <i>/-tv-/</i>	<i>/tv-/</i> <i>/-tv-/</i>	[tw-] [-tw-] [-tt-]	<i>/tv-/</i> ≡ <i>/tv-/</i> ^h <i>/-tv-/</i> ≡ { <i>/-tv-/</i> ^h , <i>/-t-/</i> ^{alt} , (<i>/-p-/</i> ^{borr})} The alternate form <i>/-t-/</i> ^{alt} mostly arises in reflexes of Sanskrit absolutes with linking <i>-i-</i> , while <i>/-p-/</i> ^{borr} is possibly a borrowed form.
<i>/dv-/</i>	<i>/dv-/</i>	[dw-] [dvv-]	<i>/dv/</i> ≡ { <i>/dv/</i> ^h , <i>/dvv/</i> ^{tr} , <i>/du-/</i> ^{tr} , (<i>/b-/</i> ^{borr}), (<i>/d-/</i> ^{borr})} It is assumed that the historic form <i>/dv/</i> ^h is retained, although <i>/dvv/</i> ^{tr} and <i>/du-/</i> ^{tr} are possibly transitional allophones. <i>/d-/</i> ^{borr} and <i>/b-/</i> ^{borr} are assumed to be borrowed forms.

8.3.1.4.3 Stop + *y* clusters

Examples: Stop + *y* clusters are very common in OIA, especially in medial positions. Gāndhārī reflexes of stop + *y* clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.1.4.3a.

Table 8.3.1.4.3a: Gāndhārī reflexes of stop + *y* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>-ky-</i>	<i>-ky-</i>	InsK-M	<i>Śākyamuṇi-</i> > <i>Śākyamuṇi-</i> (#) (1)-Wardak vase	2 CE
	<i>-k(k)-</i>	Aś-H	<i>śakyam</i> > <i>śako</i>	3 BCE
		InsK-M	<i>Śākyamuṇi-</i> > <i>Śakamuṇi-</i>	1 BCE

		InsK-M	<i>Śākyamuṇi-</i> > <i>Śakamuṇi-</i>	1 CE
		BLS1-1	<i>śākyasimha-</i> > <i>śakasiho-</i>	1 CE
		RS-14	<i>śākyamuṇi-</i> > <i>śakamuṇi-</i>	2 CE
-khy-	-(k)kh-	Aś-H	<i>mukhya-</i> > <i>mukha-</i> , <i>saṃkhyāya</i> > <i>saṃkhyay[a]</i>	3 BCE
		BLS1-2	<i>ākhyātaḥ</i> > <i>akhade</i>	1 CE
		BLS4-1	<i>saṃkhyā-</i> > <i>sakha-</i>	
		RS-19	<i>pratyākhyāya</i> > <i>pacakhae</i>	2 CE
-gy-	-g-	InsK-M	<i>ārogya-</i> > <i>aroga-</i>	1 CE
		InsK-M	<i>ārogya-</i> > <i>aroga-</i> (1)	2 CE
		BCS5-1	<i>ārogya-</i> > <i>aroga-</i>	2 CE
cy-	c-	BLS1-1	<i>cyāvayet</i> > <i>cavae</i> , <i>cyutvā</i> > <i>cavitvaṇa</i> , <i>cyutaḥ</i> > <i>cudu</i>	1 CE
		BLS4-1	<i>cyuta-</i> > <i>cudo-</i>	
-cy-	-c(c)-	InsK-M	<i>prācyah</i> > <i>pracu</i>	1 BCE
		InsK-M	<i>prācyah</i> > <i>pracu</i>	1 CE
		BLS1-2	<i>ucyate</i> > <i>bucadi</i>	1 CE
		BLS4-1	<i>ucyate</i> > <i>vucadi</i>	
		BCS5-1	<i>ucyate</i> > <i>vucadi</i>	2 CE
		RS-5	<i>ucyate</i> > <i>vucadi</i>	2 CE
	-j(j)-	RS-5	<i>ucyate</i> > <i>vujadi</i> (#) -possibly orthographic	2 CE
-chy-	-(c)ch-	Aś-H	<i>uchyati</i> > <i>vuchati</i>	3 BCE
jy-	j-	Aś-H	<i>jyotiṣ-</i> > <i>joti-</i>	3 BCE
		InsK-M	<i>jyeṣṭha-</i> > <i>jeṭha-</i>	1 CE
		InsK-M	<i>jyeṣṭha-</i> > <i>jeṭha-</i>	2 CE
-jy-	-j(j)-	InsK-M	<i>rājya-</i> > <i>raja-</i>	1 CE
		BLS1-1	<i>pravrajyā-</i> > <i>prabaja-</i>	1 CE
		InsK-M	<i>rājya-</i> > <i>raja-</i>	2 CE
		BCS5-1	<i>bhaiṣajya-</i> > <i>bheṣaje-</i>	2 CE
		RS-20	BHS <i>sajyoti-</i> > <i>sajedi-</i>	2 CE
	-(y)y-	BLS2-2	<i>udyujyate</i> > <i>ujayadi</i> (#) - uncertain	1 CE
-ṇḍy-	-ṇḍ-	Aś-H	<i>Pāṇḍya-</i> > <i>Pamda-</i>	3 BCE
	-ṇḍiy-	Aś-H	<i>Pāṇḍya-</i> > <i>Pamdiya-</i> (#)	
-ḍhy-	-(d)dh-	InsK-M	<i>ādhyā-</i> > <i>adha-</i>	1 CE
ty-	c-	BLS4-1	<i>tyāga-</i> > <i>caga-</i>	1 CE

<i>-ty-</i>	<i>-ty-</i> <i>-tiy-</i> <i>-c(c)-</i>	InsK-M	<i>parityāga-</i> > <i>parityaga-</i> (#)	3 CE
		InsK-M	<i>caitya-</i> > <i>cetiya-</i> , <i>pratyamśa-</i> > <i>pratiyamśa-</i> (#)	3 CE
		Aś-H	<i>apatya-</i> > <i>apaca-</i>	3 BCE
		BLS1-1	<i>catuḥ-satya-</i> > <i>cadu-saca-</i> , <i>pratya-</i> > <i>praca-</i> , <i>satkr̥tya-</i> > [sa]kacu-	1 CE
		BLS2-1	<i>pratyekabuddha-</i> > <i>pracegabudha-</i>	
		BLS2-2	<i>anityatā-</i> > <i>anicada-</i>	
		BLS4-1	<i>anityāḥ</i> > <i>aṇica</i> , <i>jātyandhaḥ</i> > <i>jacadho</i>	
		BLS9-1	<i>nityam</i> > <i>ṇico</i>	
		BCS5-1	<i>atyanta-</i> > [a]cata-, <i>pratyarha-</i> > <i>picara-</i>	2 CE
		RS-5	<i>yathāpratyaya-</i> > <i>yaṣapacea-</i>	2 CE
		RS-14	<i>praṇipatya-</i> > (*pra)ṇavaca-	
		RS-19	<i>pratyākhyāya</i> > <i>pacakhae</i>	
		RS-20	<i>anityāḥ</i> > <i>aṇica</i>	
		RS-20	<i>ārya-satyah</i> > <i>aria-ṣaja</i> (c > j) (#)	2 CE
<i>-thy-</i>	<i>-thy-</i> <i>-(c)ch-</i>	InsK-M	<i>mithyā-</i> > <i>mithya-</i> (#) (1)	2 CE
		BLS1-1	<i>mithyā-</i> > <i>micha-</i>	1 CE
		BLS4-1	<i>yathātathyam</i> > <i>yasatacha</i>	
		RS-19	<i>mithyā-</i> > <i>micha-</i>	2 CE
<i>-dy-</i>	<i>-(y)y-</i>	Aś-H	<i>udyāna-</i> > <i>uyana-</i> (#)	3 BCE
		BLS1-1	<i>udyānabhūmim</i> > <i>uy(y)aṇa[h]omi</i> (#)	1 CE
	<i>-j(j)-</i>	Aś-H	<i>adya-</i> > <i>aja-</i> , <i>pratipadyate</i> > <i>paṭipajeya</i>	3 BCE
		BLS1-1	<i>khādyamānam-</i> > <i>khajatu-</i>	1 CE
		BLS1-2	<i>adyāgreṇa</i> > <i>ajavaghreṇa</i> , <i>āpadyate</i> > <i>avajadi</i>	
		BLS2-2	<i>adya-</i> > <i>aj[o]-</i> , <i>udyujyate</i> > <i>ujayadi</i>	
		BLS4-1	<i>avidya-</i> > <i>avija-</i> , <i>anavadya-</i> > <i>aṇavaja-</i>	
		BLS9-1	<i>adya-</i> > <i>aja-</i> , <i>vyapanudya-</i> > <i>vavaṇuj-a</i>	
		BCS5-1	<i>vidyate</i> > <i>vijadi</i>	2 CE
		RS-5	<i>utpadyeta</i> > [u]pajea	2 CE
		RS-19	<i>āpadyate</i> > <i>avajadi</i>	
		RS-20	<i>adyāgreṇa</i> > <i>ajavagreṇa</i>	
	<i>-c(c)-</i>	BLS2-2	<i>adya-</i> > <i>aca-</i> (several) (j > c) (#)	1 CE
		RS-12	<i>niṣadya</i> > <i>niṣaca</i>	2 CE
		RS-20	<i>utpadyate</i> > <i>apacadi</i>	

<i>dhy-</i>	<i>j-</i>	BLS2-2	<i>dhyānam</i> > <i>jano</i>	1 CE
		RS-5	<i>dhyānānām</i> > <i>jaṇaṇa</i>	2 CE
		RS-12	<i>dhyānam</i> > <i>jaṇa</i>	
	<i>ḡ-</i>	BLS9-1	<i>dhyānam</i> > <i>ḡano</i>	1 CE
<i>-dhy-</i>	<i>-j(j)-</i>	BLS1-1	<i>madhyamakam</i> > <i>majimao</i>	1 CE
		BLS1-2	<i>abhidhyā-</i> > <i>abhiḡa-</i>	
		BLS4-1	<i>adhyātmam</i> > <i>ajatva</i>	
		BCS5-1	<i>madhya-</i> > <i>maja/e-</i>	2 CE
		BCS5-1	<i>adhyātmika-</i> > <i>ajatvia-</i>	2 CE
		RS-5	<i>bodhyaṅgānām</i> > <i>bejagaṇa</i>	2 CE
		RS-19	<i>-madhye</i> > <i>-maja</i>	
		RS-20	<i>adhyātmam</i> > <i>ajatva</i> , <i>acatva</i> (<i>j</i> > <i>c</i>)	
	<i>-ḡ(ḡ)-</i>	BLS1-1	* <i>vibudhyitvā</i> > <i>vioḡita</i>	1 CE
		BLS4-1	<i>adhyātma-</i> > <i>aḡatva-</i>	
		BLS9-1	<i>-madhye</i> > <i>-maḡe</i>	
<i>-py-</i>	<i>-py-</i>	InsK-M	<i>arūpyatā-</i> > <i>arupyata-</i> (#) (1)	2 CE
	<i>-p(p)-</i>	BLS4-1	<i>lipyate</i> > <i>lipadi</i>	1 CE
		BCS5-1	<i>ārūpyadhātu-</i> > <i>arupadhadu</i>	2 CE
<i>-bhy-</i>	<i>-(b)bh-</i>	Aś-H	<i>-ibhyeṣu</i> > <i>-ibheṣu</i> (Woolner 1924: 117)	3 BCE
		InsK-M	<i>abhyutsava-</i> > <i>abhusava-</i>	1 CE
		RS-5	<i>abhyavakāśa-</i> > <i>abhōaśa-</i>	2 CE

Other OIA combinations not recorded in the Gāndhārī source materials are *ghy*, *chy*, *ṭy*, *ṭhy* and *by*.

Overview: The majority of the OIA stop + *y* clusters are represented in the Gāndhārī reflexes. The number of word initial cases is limited however, with only a small number of cases recorded for OIA *cy-*, *jy-*, *ty-* and *dhy-*. These are all reduced to the single stop only, which is palatalised to *c-* or *j-* in the case of the OIA the dental stops. Medial reflexes are typically assimilated into geminate stops, recorded as single stops in the Kharoṣṭhī script (§ 8.3.1.1), while the dental stops are also palatalised as *-c(c)-* or *-j(j)-*, with use of the diacritic marked reflex *ḡ* by some of the BL scribes assumed to indicate gemination (§ 5.3.1). Some variant reflexes are also recorded, including retention of the original cluster, insertion of epenthetic *i*, and weakening to *y* only, although these are atypical. In the few instances where postconsonantal *y* is retained it is represented in the Kharoṣṭhī script by

various diacritic marks attached to the base of the consonantal akṣara, most commonly a wavy line. These include *bhye* (𑀧𑀺𑀭𑀸𑀓) in the Aśokan inscriptions, *tya* (𑀧𑀺𑀭𑀸𑀓𑀲) from the Schoyen manuscripts (Glass 2000: 119-121), while further examples of *kya*, *pya*, and *thya* from the Wardak vase are shown in § 5.3.2.1. The cases in which these occur are identified in the above table, while the example of *ārogya* > *aroga* from the same vase indicates the use of variable forms within the same inscription.

Discussion: Changes to the stop + *y* clusters indicate deletion of the *y* in word initial positions, and full assimilation of *y* by the preceding stop in medial positions, along with concurrent palatalisation of dental stops. These reflexes are similar to the MIA developments (§ 9.1.2). The semi-vowel *y* is one of the weakest coronal consonants, more prone to lenition and deletion than the other semi-vowels (Gordon 2016: 153-154, 157), and its loss in the articulatory transition between the stop and following vowel in word initial positions is consistent with this assessment. On the other hand, and similar to the other semi-vowels, it is expected that *y* would have a robust perceptual identity due to its well-defined formant structure, making it less susceptible to assimilation by the preceding stop in medial clusters. Hence although the progressive assimilation of *y* is similar to the MIA development, the other OIA semi-vowels are typically retained in Gāndhārī TR clusters (§§ 8.3.1.4.1, 8.3.1.4.2, 8.3.1.4.4). This raises doubts about whether this atypical development represents an actual sound change, or could have possibly resulted from borrowings, orthographic variation or even orthographic laxity. Although borrowed forms are a common occurrence in the reflexes of many clusters, and appear to be a natural part of the transcription process between historic documents in different dialects, it seems unusual that such borrowing would become so widespread without a foundation in the spoken language. It also seems unusual that similar borrowed forms would not also be more widely adopted for the other semi-vowels.

Historic spellings were retained in some of the Kharoṣṭhī inscriptions, seen in the examples of Skt. *Śākyamuṇi*- > G *Śākyamuṇi*-, *mithyā*- > *mithya*- and *arūpyatā* > *arupyata*- from the Wardak vase, and *parityāga*- > *parityaga*- in the Tor Ḍherai potsherds (Mehendale 1948: 308, § 522), although this usage was not widespread. Gāndhārī *Śākyamuṇi* appears to be a Sanskritised form (Konow 1929: cvi), pointing toward the use of conservative orthography in these cases rather than the recording of a general phonetic retention of the original sound. The insertion of epenthetic *i*, seen in Skt. *caitya* > G *cetiya*- in the Jamālgaṛhi pilaster base, and *pratyamśa*- > *pratiyaṃśa*- in the Tor Ḍherai

potsherds (Mehendale 1948: 308, § 522), is more interesting in this sense. Since the phonetic and acoustic nature of the semi-vowel *y* and vowel *i* are very similar (Kent and Read 2002: 177-179), the insertion of epenthetic *i* can be viewed as an allophonic transcription of the retained cluster. This suggests that the original pronunciation may have been retained in some instances at least, although the extent of this evidence is very limited. The case for orthographic laxity relates to the more complex written form of the post-consonantal *y* diacritic, which requires the addition of a downward wavy line beneath the main character, compared with the easier upward flourish and horizontal bar used for the postconsonantal *v* and *r* diacritics respectively (§ 5.3.2.1). And although the example of *ārogya* > *aroga* from the Wardak vase demonstrates a variable use of the diacritic, this interpretation assumes that such lax orthographic practices became widespread among the Buddhist manuscript scribes, which seems unlikely.

In the absence of a convincing case for these alternative interpretations, it is likely that the orthographic record represents an actual sound change that occurred in Gāndhārī, which is also supported by the similar -Cy- > -CC- developments recorded for both the Ny and Sy clusters (§§ 8.3.2.4.3, 8.3.3.4.3). All these cases entail apparent progressive assimilation of the second consonant in clusters having a profile of reducing strength, and although similar to the wider treatment of all of the C + semi-vowel clusters in MIA (§ 9.2.2), it is atypical among these types of clusters in Gāndhārī. It is possible these reflexes may represent an early progression of this development in Gāndhārī, and articulatory differences between *y* and the other semi-vowels suggest that *y* may be the most susceptible to such changes. As noted previously *y* has an open and vowel like articulation, and although *r* is considered weakest of the OIA semi-vowels in the MIA and Pali strength-based assimilation hierarchy (§ 6.1.3), *y* [j] is seen as the weakest of the coronal semi-vowels in the classification presented by Gordon (2016: 153), more prone to lenition and deletion compared with rhotics and laterals. In this context it is possible that the cluster was reduced to a single stop only following deletion of the semi-vowel. On the other hand, the normal interpretation of unweakened single stop reflexes is that they represent geminates, pointing towards progressive manner assimilation similar to the typical MIA development. One suggestion is that this apparent ‘assimilation’ could represent the diachronic development of anticipatory consonant ‘doubling’ across the syllable boundary, which is recorded as allophonic tendencies in the early Sanskrit accounts of doubling, and is discussed further in § 9.3. This appears to have been

concurrent with the lenition of *y*, as in *-Ty-* > *-_TTy-* > *-T.T^y-* > *-T.T-* for example. Most examples listed in Table 8.3.1.4.3a involve palatal or palatalised dental stops, which were possibly pronounced with a degree of affrication (§ 8.2.1.2.2), and it is possible that a phonetic trace of unrecorded *y* is also present as weakly articulated *ya-śruti*.

The alternative development of *dy* > (*y*)*y* is recorded in two cases involving addition of the verbal prefix *ud-* to a word beginning in *y*, as in Skt. *ud-yāna* > G *u(y)yana* (Aś-H) and *ud-yānabhūmim* > *u(y)yaṇa[h]omi* (BLS1-1). These reflect the same pattern as Pali and the other MIA languages (Salomon 2008: 119), suggesting the possibility of borrowings in these cases. Torwali

Summary: a summary of the Gāndhārī reflexes of OIA stop + *y* clusters is presented in Table 8.3.1.4.3b. The recorded forms are advanced in the majority of cases, and although historic forms and some variants are also represented, it is assumed that the orthographic record represents an actual sound change that occurred or was occurring in Gāndhārī. While the reasons for this change are not altogether clear, similar to the *Ny* and *Sy* clusters described in following sections, it appears to represent an early progression of the MIA manner assimilation patterns (§§ 9.2.2, 9.3) in the Gāndhārī C + semi-vowel clusters.

Table 8.3.1.4.3b: Summary of stop + *y* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
<i>/T_dy-/</i> <i>/T_αy-/</i> <i>α ≠ d</i>	<i>/T_p-/</i> <i>/T_α-/</i>	[<i>T_p-</i>] [<i>T_α-</i>]	<i>/T_p-/</i> ≡ <i>/T_p-/adv</i> (d = dental, p = palatal) <i>/T_α-/</i> ≡ <i>/T_α-/adv</i>
<i>/-T_dy-/</i> <i>/-T_αy-/</i> <i>α ≠ d</i>	<i>/-TT_p-/</i> <i>/-TT_α-/</i>	[<i>-TT_p-</i>] [<i>-TT_α-</i>]	<i>/-TT_p-/</i> ≡ { <i>/-T_p-/adv</i> , (<i>/y/alt</i>)} <i>/-TT_α-/</i> ≡ { <i>/-T_α-/adv</i> , (<i>/-T_αy-/h</i>)} Historic allographs <i>/Ty/h</i> are used occasionally, while <i>/y/alt</i> records the reflex of # <i>ud</i> + <i>y</i> clusters.

8.3.1.4.4 Stop + *r* clusters

Examples: Stop + *r* clusters are very common in OIA, in both initial and medial positions. Gāndhārī reflexes of OIA stop + *r* clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.1.4.4a.

Table 8.3.1.4.4a: Gāndhārī reflexes of stop + *r* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>kr-</i> <i>-kr-</i>	<i>kr-</i>	BLS1-1	<i>kṛta-</i> > <i>krida-</i> / <i>kridu-</i> (OIA <i>kṛ</i>)	1 CE
		BLS2-1	<i>kṛṣṇaḥ</i> > <i>kriṣo</i> (OIA <i>kṛ</i>)	
		BLS4-1	<i>krodha-</i> > <i>krosa-</i>	
		BLS9-1	<i>krīḍām</i> > <i>kriḍa</i>	
	<i>-kr-</i>	Aś-H	<i>atīkrāntam</i> > <i>atīkratam</i>	3 BCE
		BLS1-2	<i>cakrāṇi</i> > <i>cakra</i> , <i>śukraḥ</i> > <i>śukro</i>	1 CE
		BLS2-1	<i>śakraḥ</i> > <i>śakro</i>	
		BLS4-1	<i>ākruṣṭaḥ</i> > <i>akroṭha-</i> , <i>atīkramam</i> > <i>adīkramo</i>	
		BLS9-1	<i>saṃkrīḍanā-</i> > <i>saṃkriḍaṇa-</i>	
		RS-12	<i>-kram-</i> > <i>-kram-</i> (common)	2 CE
		RS-20	<i>niṣkramanti</i> > <i>ṇakramati</i>	2 CE
		InsK-M	<i>anukrama-</i> > <i>anukrama-</i>	2 CE
		BLS1-2	<i>abhīkrāntam</i> > <i>abhīkatu</i> (#)	1 CE
		RS-12	<i>prākramūt</i> > <i>pakama</i> (scribal error)	2 CE
	<i>-k(k)-</i>			
<i>gr-</i> <i>-gr-</i>	<i>gr-</i>	BLS4-1	<i>granthaḥ</i> > <i>gratha</i> , <i>grāmam</i> > <i>gramo</i>	1 CE
		BLS9-1	<i>grāme</i> > <i>gramo</i>	
		RS-19	<i>grhītvā</i> > <i>grahita</i> (OIA <i>gr</i>)	2 CE
	<i>g-</i>	BLS2-2	<i>grāmataḥ</i> > <i>gomodo</i>	1 CE
		RS-5	<i>grāmam</i> > <i>gama-</i>	2 CE
	<i>-gr-</i>	Aś-H	<i>agra-</i> > <i>agra-</i>	3 BCE
		InsK-M	<i>parigraha-</i> > <i>parigraha-</i>	1 CE
		BLS2-1	<i>-graha-</i> > <i>-graha-</i>	
		BLS4-1	<i>agra-</i> > <i>agra-</i>	
		RS-5	<i>ekāgratā-</i> > <i>ekagrada-</i>	2 CE
		RS-12	<i>samagrāḥ</i> > <i>samagra</i>	
		RS-20	<i>adyāgreṇa</i> > <i>ajavagreṇa</i>	
		InsK-M	<i>agra-</i> > <i>agra-</i>	2 CE
		InsK-M	<i>agra-</i> > <i>agra-</i> , <i>parigraha-</i> > <i>parigraha-</i> (also <i>parigaha-</i>)	3CE
		Aś-H	<i>apagrathita-</i> > <i>apaghratho-</i> (uncertain)	3 BCE
	<i>-ghr-</i>	BLS1-1	<i>agrataḥ</i> > <i>aghridu</i>	1 CE

<i>pr-</i>	<i>pr-</i>	Aś-H	<i>Priyā-darśi-</i> > <i>Priya-draśi-</i> , <i>prīti-</i> > <i>priti-</i> , <i>prāṇa-</i> > <i>praṇa-</i> , <i>pra-karaṇa</i> > <i>pra-karaṇe</i> , <i>pra-āpnoti</i> > <i>pra-punati</i> , <i>prati-bhāga-</i> > <i>prati-</i> <i>bhaga-</i>	3 BCE
		InsK-M	<i>prathama-</i> > <i>praḍhama-</i>	1 CE
		BLS1-1	<i>prati-</i> > <i>prati-</i> (many)	1 CE
		BLS1-3	<i>pra-</i> > <i>pra-</i> , <i>prati-</i> > <i>prati-</i>	
		BLS2-1	<i>prāptaḥ</i> > <i>prato</i> (many <i>pra-</i> > <i>pra-</i>)	
		BLS4-1	<i>prajña-</i> > <i>praña-</i> , <i>praṣṭhaḥ</i> > <i>praṭho</i>	
		BLS9-1	<i>pra-</i> > <i>pra-</i> (5 instances)	
		RS-5	<i>praṇihitam</i> > <i>praṇihida</i> ,	2 CE
		RS-12	<i>prajñayā</i> > <i>prañae</i>	
		RS-20	<i>prajānanti</i> > <i>prayaṇati</i>	
		InsK-M	<i>prati-graha-</i> > <i>prati-graha-</i>	
		InsK-M	<i>prapā-</i> > <i>prapa-</i>	3 CE
	<i>p-</i>	Aś-H	<i>prajā-upada-</i> > <i>pajupada-</i> , <i>pra-pautra-</i> > <i>papotra-</i> , <i>prati-padya-</i> > <i>paṭi-pajeya-</i>	3 BCE
		BLS1-1	<i>prati-</i> > <i>paḍi-</i> (common in G), <i>pra-</i> > <i>pa-</i>	1 CE
		BLS1-3	<i>prati-</i> > <i>paḍi-</i>	
		BLS4-1	<i>prati-</i> > <i>paḍi-</i>	
		BLS9-1	<i>prati</i> > <i>paḍi-</i> (4 instances)	
		BCS5-1	<i>prati-</i> > <i>paḍi-</i>	2 CE
		RS-5	<i>prati-kūla-</i> > <i>paḍi-kula-</i>	2 CE
		RS-12	<i>prati-</i> > <i>paḍi-</i> (many), <i>pravrajyām</i> > <i>pavaje</i>	
		RS-19	<i>prajānāmi</i> > <i>payañami</i> , BHS <i>prakrami</i> > <i>pakami</i>	
		Aś-H	<i>devānāmpriya-</i> > <i>devanampriya-</i>	3 BCE
<i>-pr-</i>	<i>-pr-</i>	BLS1-2	<i>abhiprasannam</i> > <i>abhiprasaṇe</i>	1 CE
		RS-12	<i>abhiprayam</i> > <i>ayiprao</i>	2 CE
		RS-5	<i>yathāpratyaya-</i> > <i>yaṣapacea-</i>	2 CE
	<i>-p(p)-</i>	RS-19	<i>apramate</i> > <i>apamate</i>	
		InsK-M	<i>abhisāraprastha-</i> > <i>Av(r)isarapatha-</i>	
	<i>br-</i>	Aś-H	<i>brāhmaṇa-</i> > <i>bramaṇa-</i>	3 BCE
		InsK-M	<i>brāhmaṇa-</i> > <i>[bra]m[h]aṇa-</i>	2 CE
		BLS1-1	<i>brahma-</i> > <i>brama-</i> (several)	1 CE

		BLS1-2	<i>brāhmaṇa-</i> > <i>bramaṇa-</i>	
		BLS4-1	<i>brāhmaṇa-</i> > <i>bramaṇo</i>	
		RS-12	<i>brahma-</i> > <i>brama-</i> (also <i>bhama</i>)	2 CE
		RS-20	<i>brāhmaṇaḥ</i> > <i>bramaṇo</i>	
<i>bhr-</i>	<i>bhr-</i>	Aś-H	<i>bhrātṛ-</i> > <i>bhrata-</i>	3 BCE
		InsK-M	<i>bhrātṛ-</i> > <i>bhratara-</i>	1 CE
		BLS1-1	<i>bhrātā</i> > <i>bhrada</i>	
		InsK-M	<i>bhrātṛ-</i> > <i>bhratara-</i>	2 CE
	<i>bh-</i>	Aś-H	<i>bhrātṛ-</i> > <i>bhatuna-</i> (Mānsehra)	3 BCE
	<i>br</i>	BLS2-2	<i>bhramareṇa</i> > <i>bramareno</i>	1 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *khr*, *ghr*, *cr*, *chr*, *jr*, *ṭr*, *ṭhr*, *ḍr* and *ḍhr*.

Overview: The OIA stop + *r* clusters are typically retained both initially and medially in Gāndhārī, although in some instances these appear to have been reduced to a single stop reflex, which is interpreted as a geminate in medial positions. Most notably the verbal prefix *prati-* is typically reduced to *paḍi-* and the prefix *pra-* is sometimes reduced to *pa-*, while other clusters are only occasionally reduced to single stop reflexes. In the Kharoṣṭhī script a stop followed by *r* is typically recorded by a horizontal diacritic underscore stroke extended to the right at the base of the *akṣara*, as in *ta* (𑀭𑀸𑀓) and *tra* (𑀭𑀸𑀓𑀲), or *pa* (𑀭𑀸𑀓) and *pra* (𑀭𑀸𑀓𑀲) for example (§ 5.3.2.1).

Discussion: Similar to the other semivowels, *r* has a well-defined formant structure giving it a robust perceptual identity (§ 6.1), and although the *y* semi-vowel was more susceptible to lenition in the Ty clusters (§ 8.3.1.4.3), the Tr clusters have remained stable in Gāndhārī. The alternate Gāndhārī reflexes of *prati-* > *paḍi-* and *pra-* > *pa-* are recorded across all of the Gāndhārī source materials, but more consistently by the RS scribe, and appear to have been copied from source texts in other MIA languages²⁴⁸. The regular use of *prati* > *prati* by the BLS1, along with the corrective marking of the *r* diacritic observed in three instances in the RS-12 manuscript (Silverlock 2015: 260), supports an interpretation that the actual pronunciation in Gāndhārī was *pr*. The occasional loss of *r* in cases such as *gr* > *g*, *tr* > *t* and *dhr* > *dh* is atypical compared with the general retention of

²⁴⁸ Salomon 2000: 89; Allon 2001: 90; Baums 2009: 170; Silverlock 2015: 259-260.

these clusters in Gāndhārī, and are most likely the result of borrowings from other MIA languages or scribal error²⁴⁹.

As discussed in § 4.1, the OIA retroflex rhotic *r* was most likely pronounced as an alveolar tap²⁵⁰. It has also been suggested, based on interpretations of the ancient treatises discussed by Baums (2009: 113), that close contact by the tip of the tongue in the tap may have been lost, leaving the root of the tongue as the primary active articulator. This might have given rise to a velar approximant such as [u], which could also be realised across a range of place options including the alveolar continuant [ɹ]. Variations such as these are seen in the different closure positions of approximant rhotics in various languages (Ladefoged and Maddieson 1996: 232-236). This suggests the possibility of articulation of *r* in the *Tr* clusters as a continuant, where the mid-cluster transition between release of the stop and contact by the tap may have weakened during opening to the following vowel.

Summary: a summary of the Gāndhārī reflexes of OIA stop + *r* clusters is presented in Table 8.3.1.4.4b. Forms are typically stable.

Table 8.3.1.4.4b: Summary of stop + *r* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/Tr-/	/Tr-/	[Tɹ-]	/Tr-/ ≡ {/Tr-/ ^h (/T-/ ^{alt})} T ≠ p
/pr-/	/pr-/	[pɹ-] [p-]	/pr-/ ≡ {/pr-/ ^h /p-/ ^{alt} }
/-Tr-/	/-Tr-/	[-Tɹ-]	/-Tr-/ ≡ /-Tr-/ ^h
			The allographs record borrowed forms from the MIA source texts, which are most common in the case of <i>pr</i> - clusters.

²⁴⁹ Baums 2009: 170; Silverlock 2015: 257

²⁵⁰ Allen 1953: 54-55; Baums 2009: 148-149.

8.3.2 Nasal plus clusters

8.3.2.1 Nasal + stop clusters

Examples: Gāndhārī reflexes of OIA nasal + stop clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.2.1a. Nasal-stop clusters only occur word medially in OIA, are typically homorganic, and are recorded using either *anusvāra* (§ 4.3, Appendix A2) or nasal + stop conjunct characters²⁵¹. Both of these approaches are used in recording the pre-consonantal nasals in Sanskrit parallels to Gāndhārī reflexes in the published source materials, and are copied in the table on that basis. There are no nasal + stop conjunct characters in the Kharoṣṭhī script, and *anusvāra* is applied inconsistently in the Gāndhārī reflexes (§ 5.3.2.3). The tabulated reflexes are again recorded as they are transliterated in the source materials.

Table 8.3.2.1a: Gāndhārī reflexes of nasal + stop medial clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>ṅk</i>	<i>k</i>	BLS4-1	<i>paṅkaḥ > pako</i>	1 CE
<i>ṁk</i>	<i>ṁk</i>	BLS9-1	<i>saṁkrīḍanā > saṁkrīḍaṇa</i>	1 CE
<i>ṁkṣ</i>	<i>kṣ</i>	BCS5-1	BHS: <i>saṁkṣiptena > sakṣiteṇa</i>	2 CE
<i>ṁkhy</i>	<i>ṁkh</i>	BCS5-1	<i>asaṁkhyāta > asaṁkheda</i>	2 CE
<i>ṅg</i>	<i>g</i>	BLS4-1	<i>kāliṅga- > kaliga-</i>	1 CE
		RS-5	<i>bodhyaṅgānām > bejagaṇa</i>	2 CE
<i>ṁg</i>	<i>ḡ</i>	RS-5	<i>astaṁgamaḥ > astaga(*mo)</i>	2 CE
<i>ṅgh,</i>	<i>ṁgh</i>	InsK-K	<i>saṅgha- > saṁgha-</i>	1 BCE
	<i>g</i>	RS-5	<i>siṅghāṇakaḥ > śigaṇa</i>	2 CE
<i>ṁgh</i>	<i>gh</i>	BLS1-2	<i>-saṁgham > -sagha</i>	1 CE
		BCS5-1	<i>saṁgham > sagha</i>	2 CE
	<i>g</i>	RS-20	<i>-saṁgham > -saga</i>	2 CE
<i>ṅc</i>	<i>ṁc</i>	Aś-M	<i>pañca- > paṁca-</i>	3 BCE
		InsK-M	<i>pañca- > paṁca-</i>	1 BCE
		InsK-M	<i>pañca- > paṁca-</i>	1 CE
		InsK-M	<i>pañca- > paṁca-</i>	3 CE

²⁵¹ Many combination characters representing Sanskrit NC clusters are used in the Devanāgarī script. Refer to the listing of compound consonant characters in Macdonell (1927: 6-8), for example.

<i>ṃc</i>	<i>c</i>	BLS4-1	* <i>ākṃcanya-</i> > <i>akicaña-</i>	1 CE
		RS-5	<i>abhimuñcati</i> > <i>avimucadi</i>	2 CE
	<i>ṃc</i>	BCS5-1	<i>-pañca-</i> > <i>-paṃca-</i>	2 CE
	<i>c</i>	BLS4-1	<i>ṃcana</i> > <i>kicano</i>	1 CE
<i>ñj</i>	<i>ṃñ</i>	Aś-H	<i>vyañjanataḥ</i> > <i>vaṃñanate</i>	3 BCE
	<i>ñ</i>	Aś-M	<i>vyañjana-</i> > <i>vañana</i>	
	<i>ṃj</i>	InsK-M	G: <i>Mu[ṃ]javaṃda-</i>	1 CE
	<i>j</i>	BLS1-2	<i>anuvyañjana-</i> > <i>aṇovejana-</i>	1 CE
<i>ṇṭh</i>	<i>ṭ</i>	RS-5	<i>utkaṇṭhati</i> > <i>ukaṭaḍi</i>	2 CE
<i>ṇḍ</i>	<i>ṃḍ</i>	InsK-M	<i>aṇḍaja-</i> > <i>a[ṃ]ḍaja-</i>	2 CE
		InsK-M	<i>Khaṇḍavana-</i> > <i>Khaṃḍavana-</i>	3 CE
	<i>ḍ</i>	BLS1-2	<i>paṇḍitena</i> > <i>paḍidena</i>	1 CE
		BLS4-1	<i>piṇḍapāta-</i> > <i>piḍavada-</i>	
		RS-5	<i>aṇḍakāni</i> > <i>aḍagaṇ[i]</i>	2 CE
		RS-20	<i>paṇḍitaḥ</i> > <i>paḍida</i>	
		InsK-M	<i>daṇḍa-</i> > <i>daḍa-</i>	2 CE
	<i>ḍ</i>	BCS5-1	<i>paṇḍita-</i> > <i>pa[ṃ]ḍida-</i>	2 CE
<i>nt</i>	<i>ṃt t</i>	InsK-M	Mehendale 1948: p. 312: normal treatments	
		BLS1-2	<i>caḥṣuṣmantah</i> > <i>caḥṣu[a]tu, śānta-</i> > <i>śata-</i>	1 CE
	<i>t</i>	BLS4-1	<i>vadanti</i> > <i>vadati</i>	1 CE
		RS-5	<i>paryantam</i> > <i>payata</i>	2 CE
		RS-12	<i>viharanti</i> > <i>viharati</i>	
		RS-20	<i>prajānanti</i> > <i>payanati</i>	
	<i>ṃt</i>	RS-20	<i>saṃtaptah</i> > <i>sat[a]ta</i>	2 CE
<i>nth</i>	<i>th</i>	BLS4-1	<i>grantha-</i> > <i>gratha-</i>	1 CE
<i>nd</i>	<i>ṃḍ d</i>	InsK-M	Mehendale 1948: p.312: normal treatments	
	<i>d</i>	RS-20	<i>abhinanditaḥ</i> > <i>aviṇaḍiḍa</i>	2 CE
	<i>(ṃ)d</i>	RS-12	<i>nandikaḥ</i> > <i>ṇaḍio</i>	2 CE
	<i>d</i>	BCS5-1	<i>sāṇḍṛṣṭika-</i> > <i>sadriṭhia-</i>	2 CE
	<i>(ṃ)d</i>	RS-12	<i>saṃdarśayati</i> > <i>sadaṣeṣi</i>	2 CE
<i>ndh</i>	<i>ṃdh dh</i>	InsK-M	Mehendale 1948: p.312: normal treatments	
		BLS1-2	<i>gandharva-</i> > <i>ghadharva-</i>	1 CE
	<i>dh</i>	BLS4-1	<i>abhandhanaḥ</i> > <i>a[va]dhaṇo</i>	
	<i>(ṃ)dh</i>	RS-12	BHS <i>anubandhitvā</i> > <i>aṇuvadhita</i>	2 CE

<i>mp</i>	<i>(m)p</i>	RS-12	<i>upasampadya</i> > <i>vasapaca-</i>	2 CE
<i>mpy</i>	<i>p</i>	BLS4-1	* <i>avikampyaḥ</i> > <i>avigapio</i>	1 CE
<i>mp</i>	<i>p</i>	BLS1-2	<i>saṃpadayati</i> > <i>sapadedi</i>	1 CE
		BLS4-1	*- <i>saṃpannaḥ</i> > - <i>sapaṇa</i>	1 CE
		BCS5-1	<i>saṃpatti</i> > <i>sapati</i>	2 CE
<i>mb</i>	<i>mb</i>	InsK-M	<i>kuṭumbinī</i> > <i>kuṭim̐bini-</i>	2 CE
		InsK-M	<i>ambā-</i> > <i>aṃba-</i>	3 CE
	<i>b</i>	InsK-M	<i>kuṭumbinī</i> > <i>kuṭibini-</i>	1 CE
		BLS1-2	BHS - <i>saṃbodhyaṅgam</i> > - <i>sabujaghu</i>	1 CE
		BLS4-1	<i>ambujaḥ</i> > <i>abuya-</i>	1 CE
<i>mbh</i>	<i>(m)bh</i>	RS-12	<i>kimbhiraḥ</i> > <i>kibhire</i>	2 CE
	<i>mb</i>	BLS9-1	<i>acchambhī-</i> > <i>acham̐bi-</i>	1 CE
<i>mbh</i>	<i>bh</i>	BLS4-1	<i>saṃbhavati</i> > <i>sabhodi</i>	1 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *ṅkh*, *ṇṭh* and *ṇdh*, but otherwise all nasal-stop clusters are well represented.

Overview: The most typical development of the OIA nasal + stop clusters in Gāndhārī involves retention of the unweakened stop. This is supplemented by the marking of *anusvāra* on the preceding syllable in some cases, as in *a* (𑀅) and *aṃ* (𑀅𑀭), and *ta* (𑀭𑀅) and *taṃ* (𑀭𑀅𑀭) for example (§ 5.3.2.3), indicating post-vocalic nasalisation at the syllable boundary and retention of the original NT cluster. The presence of unweakened stop reflexes²⁵² supports a similar interpretation in those cases where a preceding *anusvāra* is not marked, especially in locations where an NT cluster is expected through OIA parallels. There are no conjunct nasal-stop *akṣaras* in the Kharoṣṭhī script, and it appears the Gāndhārī scribes were comfortable with the irregular use of *anusvāra* and contextual interpretation of unweakened stop reflexes. Hence both *saṃgha-* or *sagha-* are recorded, while weakened forms such as **saha-*, which would only be expected in reflexes of *sagha-* for example (§ 8.2.1.2.1), are never recorded in reflexes of OIA NT clusters²⁵³. The *anusvāra* diacritic appears to be employed in reflexes involving an original OIA *anusvāra*, as well as reflexes of OIA conjunct nasal-stop *akṣaras*, as in Skt. *saṃkrīḍanā* > G

²⁵² Single intervocalic stops are typically weakened in Gāndhārī (§ 8.2.1.2), and unweakened stops suggest either a geminate (§ 8.3.1.1) or nasal + stop reflex depending upon the OIA parallel.

²⁵³ Fussman 1989: 473, §33.1; Allon 2001: 88; Glass 2007: 119

saṃkriḍaṇa and *saṅgha* > *saṃgha*- for example. The diacritic is not used consistently however, as in *-saṃgham* > *-sagha* for example, and variable usage is reported in much of the inscriptional material and some manuscripts²⁵⁴, while it is mostly absent from materials prepared by BL scribes 1 and 7 and the RS scribe²⁵⁵. The atypical development of *ñj* > *mñ* recorded in the Aśokan inscriptions (Hultzsich 1925: lxxxviii) appears to be a lone exception to the general pattern, possibly in confusion with the development of *jñ* > *ññ* (§ 8.3.1.2), or alternately because palatals undergo special development.

Discussion: The absence of conjunct *akṣaras* incorporating place specific nasal symbols in the Kharoṣṭhī script meant that OIA pre-consonantal nasals retained in Gāndhārī were either marked with *anusvāra*, or if this was absent, implicit through the retention of unweakened stops. In either case homorganic nasalisation is implied, which reflects the levelling of pre-consonantal nasals and the use of *anusvāra* as a ‘generic’ symbol for nasalisation similar to its later stage usage in OIA (§ 4.3, Appendix A2).

The writing of *anusvāra* in Gāndhārī inscriptional materials continued until at least the early second century CE (Fussman 1989: 473-479, §§33.1-33.5), while the variable use of *anusvāra* continued in some Buddhist manuscripts at least until the late first or early second century CE. Although it was recorded in the Khvs-G written by BL scribe 9 for example, many other scribes including the BL scribes 1 and 7 and the RS scribe in the second century had chosen not to write *anusvāra*, rather relying upon the strength of the following stop consonant to indicate its presence. It has been suggested that the variable and declining use of *anusvāra* among Kharoṣṭhī texts may be an indication that nasalisation had become weak (Salomon 2000: 76). And while this may also suggest it was possibly even lost, the retention of single unweakened stops as reflexes of OIA stop + nasal clusters would suggest the ongoing presence of some level of nasalisation.

Homograph pairs may have arisen due to the similar convention of marking geminate stops as unweakened single stops, although these would have been pronounced differently as long as the geminate stops retained their phonetic length (§ 8.3.1.1). It is also possible that differences would have been understood contextually in the scholarly environment of the written language. At the same time, the weakening of the single intervocalic stops would have diminished the need to retain the marking of *anusvāra*. Hence while the reflexes of Sanskrit *pañca*- may be either *paṃca*- or *paca*-, the latter usage would have

²⁵⁴ E.g., Hultzsich 1925: lxxxvii; Konow 1929: civ; Fussman 1989: 473, §33.1; Salomon 2000: 76-77.

²⁵⁵ E.g., Allon 2001: 88; Baums 2009: 158-159; Glass 2007: 119.

contrasted with a word like Skt. *paca-* which was typically pronounced as *paya-* in Gāndhārī. The declining use of *anusvāra* may indicate a weakening of nasalisation in the language, and combined with the eventual shortening of the geminate stops (§ 8.3.1.1), a merger in the pronunciation of the clusters [TT, NT] > [T] would have led to the development of homonym pairs. Languages adjust for such developments in different ways, such as through lexical replacement or loss for example (Campbell 2016: 330-332), although it is difficult to assess the extent of such developments in the later periods of Gāndhārī beyond the available record.

Summary: a summary of the Gāndhārī reflexes of OIA nasal + stop clusters is presented in Table 8.3.2.1b.

Table 8.3.2.1b: Summary of nasal + stop cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-N ₁ T ₁ -/ /-ṇT ₁ -/	/-ṇT ₁ -/	[N ₁ T ₁]	/ṇT ₁ / ≡ {/T ₁ ^h , (/ṇT ₁ ^{h.alt})} Reflexes are mostly recorded using the single unweakened stop /T ₁ ^h , which represents the historic cluster, while the alternate historic form /ṇT ₁ ^{h.alt} was used until at least the second century CE.

8.3.2.2 Nasal + nasal clusters

Examples: OIA nasal + nasal clusters only occur word medially, and only two reflexes have been recorded in the Gāndhārī source materials. These involve Skt. *nm* > G *m* or (*m*)*m*, seen in *yanmātra* > *yamatra* in the third century BCE Aśokan inscriptions (Hultsch 1925: lxxxviii) and *tanmayah* > *ta(m)maya* in the first century CE Nid-G (Baums 2009: 157, 161, 643), along with the geminate *-nn-* > *-ṇ(ṇ)-* recorded in the second century CE Gos-G (Silverlock 2015: 249)²⁵⁶.

Other OIA combinations not recorded in the Gāndhārī source documents are *ṇm*, *ṇn*, *ṇṇ*; *ṇm*, *ṇṇ*, *mm*, *mn* and *mṇ*.

²⁵⁶ These examples are again recorded as written in the source documents.

Overview: Based on the limited number of examples, it appears that in clusters of dissimilar nasals the first element assimilates to the place of the second, and the cluster develops as a long version of the second element. This is represented by the *akṣara* of a single nasal in the Kharoṣṭhī script. While there is an orthographic alternation of *ṇ* for *n* in documents such as the Gos-G prepared by the RS scribe (§ 8.2.2.2.2), it appears that geminate nasals were retained and recorded in the normal manner as a single *akṣara*.

Discussion: OIA clusters of dissimilar nasals form geminates through regressive place assimilation of the initial element, which is consistent with the phonetic tendencies discussed in § 6.1.2. This is further supported by the coronal plus non-coronal sequence of *-nm- > -mm-* seen in the two Gāndhārī examples, where the second labial nasal is expected to be more dominant. Although the available sample is small, the reflex of *-nm- > -mm-* is recorded in both the third century BCE and first century CE, suggesting a diachronic continuity.

Summary: a summary of the Gāndhārī reflexes of OIA nasal + nasal clusters is presented in Table 8.3.2.2b.

Table 8.3.2.2b: Summary of nasal + nasal cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-N ₁ N ₂ -/	/-N ₂ N ₂ -/	[N ₂ :]	/N ₂ N ₂ / ≡ /N ₂ / ^{adv}

8.3.2.3 Nasal + fricative clusters

Examples: Nasal + fricative clusters only occur in word medial positions in OIA.

Examples of their occurrence include: post-vocalic *anusvāra* preceding sibilants, as in nominals such as *vaṃśa-* and *siṃha-*; at morpheme boundaries following the preverb *sam-*, as in *sam-skṛta* (√*kr*) and *sam-sāra* (√*sr*); declensional endings such as *-āṃsaḥ* and *-āṃsi*; and in external sandhi combinations between final *-n* and voiceless coronal stops, as in *tasmin + ca = tasmīṃśca* and *narān + tatra = narāṃstatra*. Reflexes of three OIA clusters have been recorded in Gāndhārī, which are presented in Table 8.3.2.3a.

Table 8.3.2.3a: Gāndhārī reflexes of nasal + fricative clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-ṃs-	-ṃs-	BCS5-1	<i>saṃsāra-</i> > <i>saṃsa[ra]-</i> (pp. 81, 295, 309, 310)	2 CE
	-s-	BLS4-1	<i>saṃsāram</i> > <i>sasaro</i> (pp. 55, 682)	1 CE
	-ṃś-	InsK-M	<i>saṃsāra-</i> > <i>saṃśara-</i>	2 CE
	-ś-	BCS5-1	BHS <i>ānuśaṃsa-</i> > <i>aṇuśaśa-</i>	2 CE
	-ś-	BCS5-1	√ <i>śaṃs</i> > √ <i>śaś-</i> (p. 309)	2 CE
-ṃś-	-ṃś-	InsK-M	<i>pratyāṃśa-</i> > <i>pracaṃśa-</i> (p. 314)	2 CE
	-ś-	InsK-M	<i>viṃśa-</i> > <i>viśa-</i>	2 CE
	-ś-	BCS5-1	<i>viṃśati</i> > <i>viśadi</i>	2 CE
-ṃh-	-h-	InsK-M	<i>Siṃhamitra-</i> > <i>Sihamitra-</i>	3 CE

An OIA combination that is not recorded in the Gāndhārī source materials is *ns*.

Overview: Based on the limited record it appears that OIA nasal plus fricative clusters remain unchanged, but with an inconsistent recording of *anusvāra* similar to that observed for nasal plus stop clusters (§ 8.3.2.1). At the same time there has been a variable use of the palatal sibilant *ś* in reflexes of the *-ṃs-* cluster. Although represented by the generic *anusvāra*, it is possible that the preceding nasals in these OIA clusters could in some instances have assimilated to the place of the following fricative.

Discussion: Fricatives display strong internal frication which provides reliable cues to their manner and place identity, while nasals are more prone to place assimilation and weakness (§6.1.2). Like the pattern seen for the nasal + stop clusters, the non-writing of *anusvāra* again suggest that nasalization had become relatively weak or was even lost.

The variant forms recorded in the development of *-ṃs-* > {-ṃs-, -s- (-ṃś-, -ś-, -ś-)} may be interpreted in several ways. For example, the reflex of BHS *ānuśaṃsa* > G *aṇuśaśa*, which is only recorded in one document, possibly reflects palatalization or partial assimilation in response to the prior palatal sibilant, while the single example of Skt. *saṃsāra* > G *saṃśara* from the second century InsK contrasts with *saṃsa[ra]* from BCS5-1 in the same era. It is also possible that the recording of palatal sibilants in some words is reflective of the source documents, as for example in some eastern dialects where all of the sibilants merged to the palatal fricative *ś* (Misra 1967: 124). The writing of an

underscore \underline{s} in Skt. *viṃśati* > G *viśadi* is possibly intended to mark a different pronunciation (Schlosser 2016: 42), and although this could be a misplacement of the *anusvāra* diacritic, this seems unlikely as the word occurs four times in the manuscript (p. 309). Given the very small data base it is not possible to establish any clear trends of change to the sibilants in the Gāndhārī nasal + fricative reflexes, and it must be assumed that the place features were typically preserved in these clusters.

Summary: a summary of the Gāndhārī reflexes of OIA nasal + fricative clusters is presented in Table 8.3.2.3b.

Table 8.3.2.3b: Summary of nasal + fricative cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-ṃS ₁ -/	/-ṃS ₁ -/	[N ₁ S ₁]	/ṃS ₁ / ≡ {/ṃS ₁ / ^h , /S ₁ / ^{alt} } Reflexes are equally recorded using both the historic form /ṃS ₁ / ^h , and the alternate single fricative allograph /S ₁ / ^{alt} in which <i>anusvāra</i> is not recorded.

8.3.2.4 Nasal + semi-vowel clusters

Although occurring less often than stop + semi-vowel clusters, a number of nasal + semi-vowel clusters are found in OIA, with nasal + *y* clusters being the most common.

8.3.2.4.1 Nasal + *l* clusters

Only the *ml* cluster occurs in OIA, and no reflexes of this cluster have been recorded in the Gāndhārī source materials.

8.3.2.4.2 Nasal + *v* clusters

Examples: Gāndhārī reflexes of OIA nasal + *v* clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.2.4.2a.

Table 8.3.2.4.2a: Gāndhārī reflexes of nasal + *v* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-nv-	-nuv-	InsK-M	* <i>tanvaka-</i> > <i>tanuvaka-</i>	2 CE
	-ṇ-	BLS4-1	<i>samanvāgata-</i> > <i>samuṇagada-</i>	1 CE
		RS-20	<i>samanvāgata-</i> > - <i>samuṇakada-</i>	2 CE
-ṃv-	-ṃv-	InsK-M	<i>saṃvatśara-</i> > <i>saṃvatśara-</i>	1 BCE
		InsK-M	<i>saṃvatśara-</i> > <i>saṃvatśara-</i>	1 & 2 CE
	-ṃb-	InsK-M	<i>saṃvatśara-</i> > <i>saṃbatśara-</i>	2 CE
	-b- / -v-	BLS1-2	<i>saṃvara-</i> > <i>sabara-</i> / <i>savara-</i> , <i>saṃvidyamāna-</i> > <i>savijamaṇa-</i>	1 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *ṇv* and *ṇṇv*.

Overview: Slightly mixed results are seen in the reflexes of OIA -nv- and -ṃv- clusters. While -nv- > -nuv- records the insertion of the epenthetic vowel *u* in anticipation of the following labial consonant, the case of -nv- > -ṇ- suggests unrecorded gemination. The case of -ṃv- > -ṃv-, recorded in Skt. *saṃvatśara* > G *saṃvatśara* from the first century BCE to the second century CE, suggests preservation. The alternate reflex of *sambatśara* suggests occlusion of the labial semi-vowel could alternately be an inverse spelling, while single *v* and *b* reflexes suggest laxity in the recording of *anusvāra*. Post-nasal *v* is not represented by any special diacritic in the Kharoṣṭhī script (Glass 2000: 126), nor is one required in these cases where the nasality is conveyed by *anusvāra*.

Discussion: The Gāndhārī reflexes of OIA -nv- and -ṃv- are not widely attested, and definitive assessments cannot be made from the limited material. It is suggested that the reflex of -nv- > -ṇ- seen in the Nid-G is unusual, and that *ṇ* represents preserved -nv- in this case (Baums 2009: 181). The reflex of -nv- > -nuv- can be read as an allophonic transcription (§ 10.1.1), suggesting that the original cluster could be retained. Preservation is also suggested by -ṃv- > {-ṃv-, -ṃb-}, which is consistent with the general preservation of stop + *v* clusters where the labial approximant *v* is generally stable (§ 8.3.1.4.2), although it is susceptible to occlusion in the environment of the labial nasal.

Summary: a summary of the Gāndhārī reflexes of OIA nasal + *v* clusters is presented in Table 8.3.2.4.2b. In the absence of a wider data base, it is assumed that the OIA clusters are typically preserved in the Gāndhārī reflexes.

Table 8.3.2.4.2b: Summary of nasal + *v* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-nv-/	/-ṇv-/	[nv]	/-nv-/ ≡ {/-nuv-/ ^{tr} , /-n-/ ^{alt} } The transitional form is interpreted as allophonic, while the alternate form may be a borrowing.
/-ṁv-/	/-ṁv-/	[mv]	/-ṁv-/ ≡ {/-ṁv-/ ^h , (/v-/ ^{alt} , /-mb-/ ^{tr} , /-b-/ ^{tr})} The alternate form does not record <i>anusvāra</i> . The occlusion of <i>v</i> > <i>b</i> may be allophonic or represent a transitional development.

8.3.2.4.3 Nasal + *y* clusters

Examples: Gāndhārī reflexes of OIA nasal + *y* clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.2.4.3a.

Table 8.3.2.4.3a: Gāndhārī reflexes of nasal + *y* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-ṇy-	-ñ(ñ)-	Aś-H	<i>puṇya-</i> > <i>puñā-</i>	3 BCE
		BLS1-1	<i>araṇyāt</i> > <i>arañade</i>	1 CE
		BLS2-2	<i>puṇya-</i> > <i>puñā-</i>	
		BLS4-1	<i>araṇya-</i> > <i>arañā-</i>	
		BCS5-1	<i>puṇya-</i> > <i>puñā-</i>	2 CE
	-ṁ(ñ)-	BLS9-1	<i>araṇye</i> > <i>araṁñe</i>	1 CE
		Aś-H	<i>āṇṇya-</i> > <i>anaṇṇiya-</i>	3 BCE
-ny-	-ñ(ñ)-	Aś-H	<i>anya-</i> > <i>añā-</i> , <i>manyate</i> > <i>mañati</i>	3 BCE
		BLS1-1	* <i>apramānyāḥ</i> > (* <i>pra</i>) <i>mañā</i>	1 CE
		BLS1-2	<i>anyatara-</i> > <i>añadara-</i>	
		BLS2-2	<i>anyam anyam</i> > <i>añamañā</i>	
		BLS4-1	<i>anya-</i> > <i>añā-</i>	
		RS-5	<i>śūnya-</i> > <i>śuñā-</i> , <i>manyatha-</i> > <i>mañāṣa-</i>	2 CE

development is atypical among the nasal + semi-vowel clusters which otherwise mostly remain stable in Gāndhārī (§§ 8.3.2.4.2, 8.3.2.4.4). Unlike the coronal nasals the labial nasal is more resistant to palatalisation, reflected in the development of *-my-* > *-mm-*.

It has been suggested that palatalisation of the following vowel in the compounds and derivatives of Skt. *samyak* > G *same* in Nid-G may indicate a late retention of the palatal *y* element (Baums 2009: 181), while the development of *anyatarah* > *añeare* (Glass 2007: 120) suggests a similar pattern for the *ny* cluster. The recording of *ānrñya-* > *anañiya-* and *samya-* > *samya-* in the Aśokan inscriptions, along with the retained *-my-* cluster in the Niya documents, also support the observation that the *y* sound was retained at least in part in some contexts. This interpretation is consistent with the Sanskrit gemination account of MIA assimilation (§ 9.3.4), which suggests a development such as *sa^m.myak* > *sam.m^yak* > *sam.m^ye* > *sam.me*, recorded as *same*. Nasalised approximants are not uncommon in the world's languages (Ladefoged & Maddieson 1996: 132), and mostly occur as allophones arising from the spread of nasality from nearby nasal consonants or vowels. This suggests that the straightforward progressive assimilation of *y* may also be possible, as in *sa.myak* > *sa.m^yak* > *sa.m^ma* > *sam.ma* for example, while other factors such as borrowings, scribal laxity and analogical spread may have also contributed to this atypical Gāndhārī semi-vowel development (§ 8.3.1.4.3).

Summary: a summary of the Gāndhārī reflexes of OIA nasal + *y* clusters is presented in Table 8.3.2.4.3b.

Table 8.3.2.4.3b: Summary of nasal + *y* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-N _{cor} y-/	/-ññ-/	[-ɲɲ-]	/-ññ-/ ≡ {/-ñ-/ ^{adv} , (/ -mñ-/ ^{alt})} The geminate is typically represented by a single nasal, although it is sometimes marked by <i>anusvāra</i> .
/-my-/	/-mm-/	[mm]	/-mm-/ ≡ {/-m-/ ^{adv} , (/ -mm-/ ^{alt})} The geminate is typically represented by a single nasal, although it is sometimes marked by <i>anusvāra</i> .

8.3.2.4.4 Nasal + *r* clusters



Examples: Gāndhārī reflexes of OIA nasal + *r* clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.2.4.4a.

Table 8.3.2.4.4a: Gāndhārī reflexes of nasal + *r* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-mr-	-mr-	BLS9-1	- <i>mrakṣa-</i> / - <i>mṛkṣa-</i> > -[<i>mra</i>]kṣo- (assumed)	1 CE
-mṛ-	-(<i>m</i>)r-	RS-12	<i>evaṃrūpam</i> > <i>evarua</i> (compound <i>eva</i> + <i>rūpa-</i>)	2 CE
	-ṃb-	Aś-H	<i>Tāmraparṇī-</i> > <i>Taṃbapaṃni-</i>	3 BCE

One other OIA combination that is not recorded in the Gāndhārī source materials is *nr*.

Overview: It is not possible to gain a clear sense of the Gāndhārī development of OIA *mr* from the limited number of examples, each of which is uncertain in its own way:

- The assumed case of *eva(m)rua* in RS-12 is not written as such in the manuscript (Silverlock 2015: 137 (plate 8 – line 14), 249, 251, 653), and although there is a small break in the manuscript where *anusvāra* could be written, it is unclear whether this is the case. The cluster is formed across a compound boundary in Sanskrit *evam-rūpaṃ* and is not word internal, while the Gāndhārī reflex could be a reversion to the simpler form of *eva*.
- The example of *Tambapaṇni-* in the Aśokan inscriptions is recorded with an *anusvāra* underscore marker to the preceding *ta* syllable (§ 5.3.2.3), written right to left as () (Cunningham 1877: Vol I, Plate I). And while the *aḥṣara* for *ra* () is similar to *ba*, *Tambapaṇni-* is recorded in other rock inscriptions such as Girnār and Kālasi²⁵⁷, suggesting this development was widespread in the Aśokan inscriptions and is a borrowed form in Gāndhārī.
- The example of *-mrakṣa / -mṛkṣa > [mra]kṣo-* (BLS9-1) may offer a clearer indication that the cluster is preserved, but is qualified in two ways. Firstly, this may be a reflex of the alternative Sanskrit form involving vocalic *r*, and secondly, the *aḥṣara* is written

²⁵⁷ Hultzsch 1925: 185; Mehendale 1948: 27.

on a partly damaged section of the manuscript. It is however partly legible and is confirmed by the word context (Salomon 2000: 89, 110 (Plate 1 – line 33), 232).

Although developments such as Skt. *mṛgaḥ* > G *mrugo* in the Aśokan inscriptions (Hultzsch 1925: lxxxv), and *mṛgaḥ* > *mrigo* in BLS9-1 (Salomon: 2000: 89) involve reflexes of Sanskrit vocalic *r* > *ri* / *ru*, they do offer examples of Gāndhārī *mr* clusters written in the Kharoṣṭhī script (§ 5.3.2.1). Gāndhārī *mr*- is represented in the Kharoṣṭhī script with an upper right-side bar to the *aksara* for *ma*- (𑀢), as in *mru* (𑀢𑀸) and *mri*- (𑀢𑀺) for example in the Aśokan inscriptions (Cunningham 1877: Vol I, Plate I; Glass 2000: 117). This is written with greater flourish in the hand of BL scribe 9, as in *mri* (𑀢𑀺) for example (Glass 2000: 171). In the assumed case of *[mra]kṣa* the *akṣara* for *[mra]* appears to be written as (𑀢𑀺) (pp. 117, 171).

Discussion: As discussed previously *r* has a well-defined formant structure and robust perceptual identity (§ 6.1), which contributes to the ongoing stability of *Cr* clusters in OIA and Gāndhārī. Since the perceptual cues of stops are stronger than for nasals (§ 6.1.2), *r* seems less likely to be assimilated in *Nr* clusters based on the stability of *Tr* clusters (§ 8.3.1.4.4). This is consistent with the preliminary indications provided by the above examples, and in the absence of a wider data base it is assumed that the *mr* clusters remain stable in Gāndhārī. Pronunciation of the post nasal *r* is suggested as an alveolar continuant [ɹ], for the reasons discussed in § 8.3.1.4.4.

Summary: a summary of the Gāndhārī reflexes of OIA nasal + *r* clusters is presented in Table 8.3.2.4.4b. This is based on the assumed stability of the *mr* cluster.

Table 8.3.2.4.4b: Summary of nasal + *r* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-mr-/	/-mr-/	[-mɹ-]	/-mr-/ ≡ {/-mr-/ ^h , /-ṃb-/ ^{alt} } The alternate forms are possibly borrowings.

8.3.3 Fricative plus clusters

8.3.3.1 Fricative + stop clusters

Examples: Fricative + stop clusters occur both initially and medially in OIA. Gāndhārī reflexes are recorded for eleven of the fourteen OIA clusters, and these are presented in Table 8.3.3.1a along with a selection of typical examples from the source materials. Where single unweakened stop reflexes are recorded in the Kharoṣṭhī script, they are assumed to be geminates for the reasons outlined in § 8.3.1.1. Similarly, the convention of bracketing the assimilated consonant in these geminate reflexes has been followed in the table.

Table 8.3.3.1a: Gāndhārī reflexes of fricative + stop clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-ḥkh-	-(k)kh- (h)	RS-19	<i>duḥkhite</i> > <i>dukhide</i> > <i>duhide</i>	2 CE
-śc-	-(c)ch-	Aś-H	<i>kaścit</i> > <i>kaci</i>	3 BCE
		RS-12	<i>paścāt</i> > <i>pāci</i>	2 CE
	-(c̄)ch̄- rc	BCS5-1	<i>paścāt</i> > <i>pača</i>	2 CE
		BLS1-1	<i>paścāt</i> > <i>parce</i> (#)	1 CE
-śp-	-śp-	InsK-M	> <i>Piśpas(r)i-</i> . Skt. unknown, but <i>śp</i> preserved. (Mehendale 1948: 307)	1 BCE
		BLS2-2	<i>aśpavarmā</i> > <i>[a]śpavarmo</i>	1 CE
		InsK-M	> <i>Veśpaśi-</i> . Skt. unknown, but <i>śp</i> preserved.	2 CE
-ṣk-	-ṣk-	InsK-M	<i>Kaniṣka-</i> > <i>Kaniṣka-</i> , <i>Vajheṣka-</i> > <i>Vajheṣka-</i>	2 CE
		Aś-M	<i>duṣkara-</i> > <i>dukara-</i>	3 BCE
		Aś-M	<i>duṣkṛta-</i> > <i>dukaṭa-</i>	
	-(k)kh-	InsK-M	<i>puṣkariṇī-</i> > <i>pukariṇi-</i>	1 BCE
		InsK-M	<i>puṣkariṇī-</i> > <i>pukaraṇi-</i>	1 CE
		BLS1-1	<i>niṣkaṣāyaḥ</i> > <i>ṇikaṣaya</i>	1 CE
		BLS9-1	<i>niṣkuhaḥ</i> > <i>ṇikuho</i>	
		BLS2-2	<i>puṣkalāvatī</i> > <i>pokhala[di]</i>	1 CE
		RS-5	<i>puṣkariṇīm</i> > <i>pukharana</i>	2 CE
-ṣṭh-	-(t)ṭh-	Aś-H	<i>aṣṭa-</i> > <i>aṭha-</i>	3 BCE
		Aś-M	<i>aṣṭamī-</i> > <i>aṭha-</i> , <i>śreṣṭha-</i> > <i>sreṭha-</i> , <i>tiṣṭha-</i> > <i>tiṭha-</i>	

		InsK-M	<i>aṣṭa-</i> > <i>aṭha-</i> (also <i>aḍha-</i> voiced)	1 BCE
		BLS1-1	<i>nirdiṣṭaḥ</i> > <i>ṇidiṭhu</i>	1 CE
		BLS1-2	<i>prṣṭaḥ</i> > <i>proṭhu</i>	
		BLS2-2	<i>drṣṭānaḥ</i> > <i>dr[i]ṭhate</i>	
		BLS9-1	<i>niviṣṭam</i> > <i>ṇiviṭha</i>	
		BLS4-1	<i>akṣṭaḥ</i> > <i>akroṭha</i>	
		InsK-M	<i>yaṣṭi-</i> > <i>yaṭhi-</i> , <i>aṣṭama-</i> > <i>aṭhama-</i>	2 CE
		BCS5-1	<i>drṣṭa-</i> > <i>driṭha-</i>	2 CE
		RS-5	<i>aṣṭa-</i> > <i>aṭha-</i>	2 CE
		RS-19	<i>drṣṭa-</i> > <i>driṭha-</i>	
		RS-20	<i>adrṣṭatvāt</i> > <i>adhriṭhatva</i>	
<i>-ṣṭh-</i>	<i>-(t)ṭh-</i>	Aś-H	<i>tiṣṭhati</i> > <i>tiṭhiti</i>	3 BCE
		InsK-M	<i>apraṭiṣṭhāpita-</i> > <i>apraṭiṭhāvita-</i>	1 BCE
			<i>kaniṣṭha-</i> > <i>kaniṭha-</i>	1 CE
		BLS1-1	<i>agraśreṣṭhinaḥ</i> > <i>aghrasreṭhisa</i>	1 CE
		BLS2-2	<i>suṣṭhu-</i> > <i>suṭhu-</i>	1 CE
		BLS4-1	* <i>aṇiṣṭhūrī</i> > <i>aniṭhuri</i>	1 CE
		InsK-M	<i>jyeṣṭha-</i> > <i>jeṭha-</i>	2 CE
		RS 5	<i>kāṣṭha-</i> (?) > <i>ka[ṭha]-</i>	
		BCS5-1	<i>suṣṭhu-</i> > <i>suṭhu-</i>	
		InsK-M	<i>Vāsiṣṭha</i> > <i>Vaseṭha-</i>	3 CE
	<i>-(ṭ)ṭh-</i>	BLS1-1	BHS <i>upatiṣṭhatha</i> > <i>vaṭhasa</i>	1 CE
		BCS5-1	<i>pratitiṣṭhā</i> > <i>p[aditiṭha]</i>	2 CE
	<i>-(t)ṭh-</i> <i>-st-</i>	InsK-M	<i>pratiṣṭhāpita-</i> > <i>pratithāvita-</i>	1 CE
		InsK-M	<i>pratiṣṭhāpita-</i> > <i>pratistavita-</i>	1 BCE
			<i>pratiṣṭhāpita-</i> > <i>pratistavita-</i>	1 CE
			<i>pratiṣṭhāpita-</i> > <i>pratistapita-</i>	2 CE
<i>-ṣp-</i>	<i>-śp-</i>	BLS1-1	<i>puṣṭam</i> > <i>puśpa</i>	1 CE
<i>sk-</i>	<i>kh-</i> <i>k-</i>	Aś-M	<i>skandha-</i> > <i>-khaṃda-</i>	3 BCE
		Aś-M	<i>skandha-</i> > <i>-kaṃdha-</i>	3 BCE
		InsK-M	<i>skandhāvāra</i> > <i>kadhavara</i>	1 CE
			<i>skandha-</i> > <i>-kaṃdha-</i>	2 CE
		BLS4-1	<i>skandhāḥ</i> > <i>kadha</i>	1 CE
<i>-sk-</i>	<i>-(k)k-</i>	BLS9-1	<i>-skandhaḥ</i> > <i>-kaṃdho</i>	1 CE

<i>-ṃṣk-</i>	<i>-k(k)-</i>	RS-12	<i>avaskāra-</i> > <i>avakara-</i> (Skt. <i>ava-kāra</i> ?)	2 CE
		RS-19	<i>dāruskandhaḥ</i> > <i>darukadhe</i>	
		BLS1-1	<i>puraskṛtaḥ</i> > <i>purakidu</i>	1 CE
		InsK-M	<i>saṃskāra-</i> > <i>saṃkara-</i>	2 CE
		RS-5	<i>saṃskārāḥ</i> > <i>sakhara</i> (#)	2 CE
<i>st-</i> <i>-st-</i>	<i>th-</i>	InsK-M	<i>stūpa-</i> > <i>thuba-</i>	1 BCE
		InsK-M	<i>stūpa-</i> > <i>thuva-</i>	1 CE
		BLS1-1	<i>stūpa-</i> > <i>thuba-</i>	1 CE
		InsK-M	<i>stūpa-</i> > <i>thubu-</i>	2 CE
	<i>-st-</i>	Aś-H	<i>asti</i> > <i>asti</i> , <i>hastinaḥ</i> > <i>[a]stina</i>	3 BCE
		Aś-M	<i>hastin-</i> > <i>hasti-</i>	3 BCE
		InsK-M	<i>vāstavya-</i> > <i>vastava-</i> ,	1 CE
		BLS1-1	<i>saṃstīrya</i> > <i>sastarita</i>	1 CE
		BLS2-2	<i>vistaraḥ</i> > <i>vistaro</i> and many similar,	
		BLS4-1	<i>vastu</i> > <i>vastu-</i>	
		BLS9-1	<i>-praśastam</i> > <i>-praśastam</i>	
		InsK-M	<i>daurmanastā</i> > <i>dormanasta</i>	2 CE
		RS-5	<i>mastakaḥ</i> > <i>masta</i>	2 CE
		RS-12	<i>śāstā</i> > <i>śasta</i>	
		RS-20	<i>asti</i> > <i>asti</i>	
		BLS2-2	<i>śāstā</i> > <i>śathu</i>	1 CE
	<i>-(t)th-</i>	BLS1-1	<i>sthaviraḥ</i> > <i>thero</i>	1 CE
		BLS2-2	<i>sthavikāḥ</i> > <i>thaviga</i>	
		BLS4-1	<i>sthitakāḥ</i> > <i>thidaga</i>	
		InsK-M	<i>sthūṇā-</i> > <i>thuṇa-</i>	3 CE
	<i>ṭh-</i> <i>ṭh-</i>	BLS2-2	<i>sthitaka-</i> > <i>ṭhidaga-</i>	1 CE
		BLS1-2	<i>sthitena</i> > <i>ṭhidenā</i>	1 CE
		BLS4-1	<i>sthitakāḥ</i> > <i>ṭhidaga-</i>	
		RS-12	<i>sthitā</i> > <i>aṭhīde</i>	2 CE
	<i>ṭ-</i>	BLS4-1	<i>sthitakāḥ</i> > <i>ṭida</i>	1 CE
		RS-5	<i>sthitam</i> > <i>ṭida</i>	
		RS-12	<i>sthāsyati</i> > <i>ṭasadi</i>	
		InsK-M	<i>sthitī-</i> > <i>-stithi-</i>	1 BCE
	<i>st-</i> <i>-sth-</i>	Aś-M	<i>gr̥hastha-</i> > <i>grahatha-</i> (S), <i>gahatha-</i> (M)	3 BCE

	-(<i>t</i>) <i>ṭh</i> -	InsK-M	<i>śivasthala-</i> > <i>śivathala-</i>	1 CE
		BLS1-1	<i>upasthitaḥ</i> > <i>upaṭhita</i>	1 CE
		BLS1-2	<i>asthi-</i> > <i>aṭhi-</i>	
		BLS4-1	* <i>bhūmyastham</i> > <i>bhumaṭho</i>	
		RS-5	<i>smṛtyupasthānānām</i> > <i>spadoṭhaṇaṇa</i>	2 CE
		RS-12	<i>asthāsi</i> > [<i>a</i>] <i>ṭhaṣi</i>	
	- <i>t</i> (<i>t</i>)-	BCS5-1	<i>laghūsthāna-</i> > <i>lahuṭhaṇa-</i>	2 CE
		BLS4-1	<i>parvatasthah</i> > <i>pravadaṭe</i>	1 CE
		RS-5	<i>asthi-</i> > <i>aṭi-</i>	2 CE
		RS-12	BHS <i>asthāsi</i> > <i>aṭaṣi</i>	
		RS-20	<i>asthāt</i> > <i>aṭa</i>	
	<i>st</i>	InsK-M	<i>Sakasthāna-</i> > <i>Sak(r)astana-</i>	1 CE
<i>sp-</i>	<i>ph-</i>	InsK-M	<i>sparśa-</i> > <i>phaṣa-</i>	2 CE
		BLS1-2	<i>spṛśitvā</i> > <i>phuṣita</i>	1 CE
		BLS4-1	<i>sparśita-</i> > <i>phaṣida-</i>	
		BLS9-1	<i>sparśayet</i> > <i>phaṣae</i>	
		BCS5-2	<i>spṛśati</i> > <i>phaṣati</i>	2 CE
	<i>phr-</i>	RS-12	<i>sparśā-</i> > <i>phraṣa-</i> (#)	2 CE
	- <i>śp-</i>	BCS5-1	BHS. <i>a-sprśana</i> > <i>a[śpri]ṣa[ṇaṇa]</i>	2 CE
	- <i>phr-</i>	RS-12	- <i>sparśāvihara-</i> > - <i>phraṣavihara-</i> (#)	2 CE
- <i>sp-</i>	- <i>(p)p-</i>	RS-12	- <i>sparśāvihara-</i> > - <i>paṣaviharo-</i>	2 CE
<i>sph-</i>	<i>sp-</i>	BCS5-2	<i>sphura-</i> > <i>spura-</i>	
- <i>sph-</i>	<i>ph-</i>	BLS1-1	* <i>asphaṭṭit</i> > <i>phaṭiṣ[u]</i>	

OIA combinations not recorded in the Gāndhārī source materials are *ṣkh*, *ṣph* and *skh*.

Overview: The OIA fricative-stop clusters involve combinations of the sibilant fricatives with voiceless stops only, and coronal stops only occur in homorganic clusters with the sibilant having the same place of articulation.

Word initial OIA sibilant + stop clusters are mostly reduced to an aspirated version of the stop, as in *st-* > *th-* and *sp-* > *ph-*, while the recording of *sth-* > {*th-* (*th-*, *ṭh-*, *t-*)} suggests attempts to capture irregular or changing pronunciation with diacritic marked characters in some of the reflexes. The velar case develops without aspiration, as in *sk-* > *k-*. In word medial OIA sibilant + unaspirated stop clusters a variety of reflexes are recorded:

- The palatal cluster *-śc-* changes to *-(c)c-* in Skt. *kaścit* > G *kaci* in the Aśokan inscriptions, to *-c̄-* in *paścat* > *paṁcā* by the RS scribe and BC scribe 5, and *-rc-* in *paścat* > *parce* by the BL scribe 1.
- Clusters involving the retroflex or dental sibilants with the velar stop *-{ṣ, ṣ}k-* are typically reduced to the unaspirated velar stop *-(k)k-*, although some atypical reflexes involving aspiration or diacritics are also recorded. In two second century inscriptions *-ṣk-* is preserved for evident reasons in the proper nouns *Kaniṣka* and *Vajheṣka*.
- The homorganic retroflex and dental clusters *-ṣṭ-* and *-st-* show different developments. While the retroflex cluster is always reduced to the aspirated geminate *-(t)ṭh-*, the dental cluster is mostly preserved, but appears as *-th-* in the single example of Skt. *śāstā* > G *śathu* recorded by BL scribe 2.
- Although only recorded in a limited number of sources, the sibilant plus labial stop *-{ṣ, ṣ}p-* clusters are all preserved or assimilated as *-śp-*.

Reflexes of the OIA medial sibilant + aspirated retroflex and dental stop clusters *-ṣṭh-* and *-sth-* are mostly reduced to the aspirated stops *-ṭh-* and *-th-* respectively, although these are intermixed and sometimes recorded as the diacritic-marked *-ṭ̣h-*, or as unaspirated *-t-* in one instance. These clusters are also recorded as unaspirated *-st-* in a small number of cases in the Kharoṣṭhī inscriptions. Although developments of *-ṣkh-* > *-kh-* > *-h-* and *-sph-* > *-ph-* are only recorded in single cases they appear consistent with the main pattern.

Retained and developed fricative + stop clusters are recorded using both special *akṣaras* and combined characters in the Kharoṣṭhī script (§§ 5.3.3, 5.3.4). Examples of clusters recorded in the source documents are *st* (𑀲) using a special *akṣara*, and *ś* (𑀱) + *p* (𑀧) > *śp* (𑀲𑀧) and *s* (𑀭) + *p* (𑀧) > *sp* (𑀭𑀧) which combines the individual characters. The diacritic in the *ṭh* reflex is marked by an upward extension to the left-hand arm of the character, as in (𑀲̣) (§ 5.3.1), while the form of *-c̄-* used in the writing of *paṁcā* in BC Fragment 4 employs the common horizontal superscript bar, as in (𑀭̄) (Glass 2000: 62; Schlosser 2016: 44-45).

Discussion: Sibilants have robust cues to their identity, and it might be expected that fricative plus stop clusters would have an inherent level of stability (§ 6.1). The

articulation of fricative plus stop clusters involves a period of frication during closure of the stop, which is a phonetically simple adjustment, suggesting that this is a strong position for sibilants. This is seen for example in the homorganic *st* and *ṣt* clusters in OIA, where the coronal sibilants control the place of the following stops, seen for example in the spread of retroflexion in the *-ṣt-* clusters (Appendix A1). Although ST clusters involve a sonority reversal, and these pre-stop sibilants have been classified as ‘extra syllabic’, sibilants appear to be largely free of the sonority sequencing constraints that govern other sounds (Gordon 2016: 102-103). The developments in Gāndhārī include retention in some cases, while other cases involve loss of the sibilant, which is accompanied by other processes such as gemination, or feature spreading by the sibilant²⁵⁸, in different situations. These developments vary between the stops according to their places of articulation and word position, and these patterns do not always develop in a consistent manner. The patterns of development are discussed in more detail below.

Word initial reflexes:

Reflexes of the OIA word initial unaspirated clusters *st-* and *sp-* involves deletion of the initial sibilant and concurrent aspiration of the stop in case of the *st-* > *th-* and *sp-* > *ph-* clusters. The release bursts of voiceless unaspirated stops often contain a short period of aspiration which is suppressed by a prior sibilant (Davenport and Hannahs 210: 22-23). This appears to have been realised in the Gāndhārī reflexes in a process of auditory metathesis (Gordon 2016: 172-173), as in *st* > *ts* for example, combined with debuccalisation in which the place features of the fricative spread to the glottal aspirate [h] in the post-stop position (§ 6.2), represented as *st-* > *ts-* > *th-*. One atypical reflex of *sp-* > *phr-* is recorded in Skt. *sparśā-* > G *phraṣa-* by the RS scribe in the Gos-G, which appears to reflect historical spellings in source dialects (Silverlock 2015: 268-269).

Aspiration is recorded once in the reflex of *sk-* > *kh-*, seen in Skt. *skandha-* > G *-khaṃda-* in the Aśokan inscriptions, which is probably a borrowing, but is unrecorded in the more common reflex of *sk-* > *k-*. This typically occurs in variations of *skandha-* > *(-)kadha-* / *(-)kaṃda-*, where the use of a prior dash in some cases may possibly indicate aspiration (Mehendale 1948: 22, 305). This reflex is also used by BL scribe 4 in the Nid-G

²⁵⁸ While the development such as ST > T^h can be described as one involving a combination of metathesis and debuccalisation, the presence of potential intermediate stages as in *st* > *(ht / ts)* > *th* is unclear. (E.g., Cho 1999: 183; Suzuki 2002a: 106-108). This process might also be described more simply as one of feature spreading by the sibilant.

manuscript, where its phonetic intent is most probably aligned with the diacritic marked *k'* which is regularly used by other scribes (Baums 2009: 167). It has been suggested that the diacritic marking in *k'* could reflect a trace of the original sibilant, as in *^sk* (Bailey, as cited by Salomon 2000: 91), while this might also indicate a slight aspiration of the stop in the word *saṃk'ara*, as in *-k^h* (Konow 1929: cx). Both of these suggestions are consistent with transitional stages in the development of these clusters through feature spreading by the sibilant. Combined with the actual recorded case of *sk-* > *kh-* in the Aśokan inscriptions, this suggests that an aspirated reflex was at least developing at an allophonic level if not fully developed for this cluster.

Reflexes of OIA word initial aspirated clusters are only recorded for the dental cluster *sth-* in derivations of the Sanskrit verbal root $\sqrt{sthā}$ -, mostly seen in variations of Skt. *sthV-* > G *thV-*, and involving deletion of the initial sibilant and retention of the aspiration. Several alternative reflexes are also recorded, as in *sth-* > {*th-*, *ṭh*, (*ṭh-*, *ṭ-*, *st-*)}. The simple development of *sth-* > *th-* suggests that the initial sibilant is possibly also susceptible to feature spreading in these cases, but is absorbed into the pre-existing aspiration of the stop. On the other hand, the use of alternate and diacritic marked reflexes appears to be more common in the overall record, suggesting an uncertainty about the actual phonetic nature of this development. Retroflex stops are mostly conditioned and typically occur in word medial positions in Sanskrit (Appendix A1), and the development of *ṭh-* or *ṭ-* as the reflex of the word initial dental cluster seems unlikely, suggesting the use of these graphemes to represent a different or transitional sound in this situation, which is also suggested by use of the diacritic marked reflex *ṭh*-. All of the potential reflexes have been recorded by BL scribe 4 in the Nid-G, as in Skt. *sthita-* > G *thida-*, *ṭhida-*, *ṭhida-* and *ṭida-* (Baums 2009: 163), which points towards a complex history in the scribal representation of the development of these clusters in both initial and medial positions (Baums 2009: 164-166), which could be interpreted as different degrees of analogical extension of the *ṣṭh* cluster from the present-tense *tiṣṭhati*. While it has been suggested that use of *ṭh* in Gāndhārī might be intended to represent the preserved cluster *sth-* (Brough 1962: 75-77, § 18, 18a, 18b), this sign has a complex history, and based on analysis of these various developments, Baums (2009: 166) has proposed that this be interpreted as representing phonetic *th-* or *-(t)th-*. Combined with the common *th-* reflex, this suggests a likely development of this cluster is *sth-* > *th-*, although this is offered without any great certainty due to the complex record of development for these clusters.

Word medial reflexes:

a. Clusters involving OIA medial unaspirated stops:

The homorganic dental cluster *-st-* is almost universally retained, while the *-th-* reflex recorded in Skt. *śāstā* > G *śathu* (BLS2-2) appears to be less common, and the development can be summarised as *-st-* > *{-st-, (-th-)}*. This stability contrasts with feature spreading by the sibilant seen in the word initial dental clusters, suggesting that the sibilant may be less susceptible to this process in intervocalic positions where it may form a coda in the preceding syllable (Appendix A4). Reflexes of the retroflex cluster *-ṣṭ-* are however recorded as *-ṭh-*, suggesting a process of feature spreading along with gemination of the stop. This reflex has also been interpreted to represent the preserved pronunciation, however, based on modern Dardic languages where these clusters are retained, and it is possible it was used in Gāndhārī in the absence of a conjunct retroflex character²⁵⁹. Since retroflex stops are normally conditioned by a prior retroflex sibilant in OIA, use of the *akṣara* for a single retroflex stop may have been a convenient way to represent a retained *-ṣṭ-* cluster, in a similar way to the marking of geminates by single unweakened stops, and retention makes sense in the context of the retained *-st-* clusters. The *-ḍh-* reflex appears to reflect intervocalic voicing in the case of Skt. *aṣṭa-* > G *aḍha-* recorded in the Kharoṣṭhī inscriptions (Mehendale 1948: 305), and in the absence of alternative explanations, the development at the phonemic level is suggested as *-ṣṭ-* > *{-ṣṭ-, (-ṣḍ-)}*. This interpretation is consistent with the sibilant plus labial stop clusters which develop as *-{ś, ṣ, s}p-* > *-śp-* and are effectively stable apart from the apparent place aggregation of the sibilants, although it is difficult to analyse this further due to the small sample.

The main divergence from this pattern occurs in clusters involving the velar and palatal stops, where the clusters do not appear to be retained and mostly develop as unaspirated geminate stops. The velar stop *k* occurs in OIA retroflex and dental sibilant clusters, which develop as *-ṣk-* > *{-k-, (-kh-), (-ṣk-)}* and *-sk-* > *{-k-, (-k'-)}* respectively. The assumed geminate *-(k)k-* appears to be the most common reflex, while the retained form of *-ṣk-* only occurs in inscriptional versions of proper nouns. This is instructive however since it confirms the existence of a composite conjunct character for this cluster, which is also recorded in the Dhp-G^K, as in *śa* (\mathcal{P}) + *ka* (\mathcal{K}) > *śka* (\mathcal{P}) (Glass 2000:

²⁵⁹ Brough 1962: 77 § 18b, 104 § 60; Baums 2009: 164. Such direct comparisons with the Dardic languages may not be reliable, however, due to the time distance between the two language records, as discussed in § 2.2.6 and § 8.2.6.3.

175-181), although the complexity of this character may have inhibited its regular use. The diacritic marked reflex *-ḳ-* quite possibly indicates a geminate (Salomon 2008:126), but could also indicate transitional developments such as *-^sk-* or *-^hk-*, pointing towards the fully developed aspirated reflex *-kh-*, which is also recorded. The *-kk-* and *-kkh-* reflexes are attested quite frequently in the MIA languages, and it has been suggested that this could be a junctural phenomenon which arises when the cluster was formed at OIA morpheme boundaries (Misra 1967: 143). The examples given by Misra all involve the OIA *-ṣk-* cluster in the Śāurasenī record, and the Gāndhārī reflexes of this cluster are mostly consistent with this analysis, as for example in Skt. *duṣ+kara-* > G *dukara-* (Aś-M) and *niṣ+kaṣāyaḥ* > *ṇikaṣaya* (BLS1-1) involving the verbal prefixes *du-* and *ni-*²⁶⁰, compared with *puṣkalā-vatī* > *pokhala[di]* (BLS2-2) and *puṣkariṇīm* > *pukharaṇa* (RS-5), although the example of *puṣkariṇī-* > *pukariṇi-* (InsK-M) is atypical in this context. It is interesting that none of the OIA *-sk-* clusters listed in Table 8.3.3.1a are formed across morpheme boundaries, but are similarly reduced to an unaspirated geminate, as for example in verbal cases such as Skt. *saṃ+skāra-*²⁶¹ > G *saṃkāra-* (InsK-M), *ava+skāra-* > *avakara* (RS-12) along with the compound *dāru-skandhaḥ* > *darukadhe* (RS-19), while *puras+kṛtaḥ* > *purakidu* (BLS1-1) could reflect either a slight aspiration or trace of the original sibilant as noted previously. An alternative consideration is that since none of these clusters is homorganic, the perceptual cues of the slower velar stop are more likely to overshadow the cues of the faster coronal sibilants in regressive assimilation, and the sibilant was assimilated rather than debuccalised in these cases.

Reflexes of the palatal cluster *-śc-* are recorded as the stop itself or a modified version in some way. Hence *-śc-* > *{-c-, -rc-, -c̣-}*. These are mostly contained in reflexes of OIA *kaścit*, *paścāt*, and *paścima*. It appears that both *-c̣-* and *-rc-* are intended to convey a modified pronunciation (Salomon 2008: 124), which may be either aspiration or gemination (Silverlock 2015: 265) or potentially a combination of both.

b. Clusters involving OIA medial aspirated stops:

The Gāndhārī reflexes of the OIA aspirated retroflex cluster *-ṣṭh-* are summarised as *-ṣṭh-* > *{-ṭh-, (-th-, -st-, -ṭh-)}*, in which the aspirated retroflex stop *-ṭh-* is most common, with *-th-* and *-st-* only occurring in inscriptions. As noted previously an independent

²⁶⁰ These appear as *duṣ-* and *niṣ-* in many examples where they prefix *-k-* stems (s.v. MW).

²⁶¹ The verbal root *√kr-* appears to retain the initial *s* from an older version of the root when it follows certain verbal prefixes, giving rise to derived forms such as *skāra-* (Whitney 1924: 400 § 1087d).

graphic *ṭh* may have been used to represent phonetic *-ṣṭ-*, although this may also be interpreted as *-ṣṭh-* in a reflex of the OIA aspirated cluster, suggesting a preservation of these clusters. The Gāndhārī reflexes of the aspirated dental cluster *-sth-* are summarised as *-sth-* > *{-ṭh-, -t-, (-th-, -st-)}*, in which *-th-* and *-st-* only occur in inscriptions. As noted previously, the modified graphic *ṭh* has been interpreted as representing phonetic *-(t)th-*, suggesting that feature spreading by the sibilant could have occurred in the case of the dental clusters. The reasons why the retroflex clusters may be more stable than the dental clusters, as suggested in the above analysis, is not altogether clear, although this could be related to the more complex articulatory nature of the retroflex pronunciation. The common use of both the *-ṭh-* and *-t-* reflexes for the dental clusters also suggests an attempt to capture some form of transitional pronunciation in the orthography. Although there is some intermixing of their use, the former reflex appears to have been used more frequently by the BL scribes, and the latter more so by the RS scribe, suggesting this could be a matter of local scribal conventions. Many of the words listed in the table are derivatives of the OIA verbal root *√sthā-* which include the use of both the *-ṣṭh-* and *-sth-* clusters. And while there appears to be a pattern of resolving both of these as *-ṭh-* by particular scribes²⁶², variations also occur as in Skt. *asthāsi* > G *[a]ṭhaṣi* and BHS *asthāsi* > *aṭaṣi* by the RS scribe for example, which suggests some intermixing of the scribal conventions. The OIA *-ṣkh-* and *-sph-* clusters are represented in single cases only and it is not possible to draw any meaningful assessment of their development on this basis.

Summary: a summary of the Gāndhārī reflexes of OIA fricative + stop clusters is presented in Table 8.3.3.1b. Variable reflexes are recorded in most cases, with many apparent transitional sounds or allophones represented by a series of allographs. There is however little clear indication of the actual sound which rely on interpretations in many cases.

Table 8.3.3.1b: Summary of fricative + stop cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
-------------	---------------------	---------------	---

²⁶² Salomon 2008: 125-126; Silverlock 2015: 270-271.

Word initial: the most common pattern is summarised as S{T ₁ -, T ₁ ^h -} > T ₁ ^h -			
/sk-/	/kh-/	[^s k], [k ^h]	/kh-/ ≡ {/k-/ ^{tr} , (/kh-/ ^{adv})} Transitional allophones are suggested by /k-/ ^{tr} , while these may be developing towards /kh-/ ^{adv} .
/st-/	/th-/	[t ^h]	/th-/ ≡ /th-/ ^{adv}
/sth-/	/th-/	[t ^h]	/th-/ ≡ {/th-/ ^{adv} , /t̪h/ ^{alt} , (/t̪h-/ , /t̪-/) ^{alt} } The alternate forms appear to represent transitional stages in the marking of the reflex of this cluster.
/sp-/	/ph-/	[p ^h]	/ph-/ ≡ {/ph-/ ^{adv} , (/phr-/ ^{alt})} The alternate form may relate to the source dialect.
Word medial: clusters of homorganic coronal pairs are mostly retained. Clusters involving palatal and velar stops are geminated in favour of the stop. Sp clusters are retained as śp.			
/-śc-/	/- cc-/ uncertain	[cc]	/- cc-/ ≡ {/- c̄-/ ^{adv} , (/c-/ , /rc-/) ^{alt} } The diacritics may convey either gemination or aspiration, or both.
/-ṣk-/ /-sk-/	/-kk-/ assumed	[kk]	/-kk-/ ≡ {/-k-/ ^{adv} , (/ḳ-/) ^{tr} } The stronger velar stop may have assimilated the coronal sibilants. The transitional form may indicate a geminate or a partially developed aspirated allophone.
/-ṣt̪-/	/-ṣt̪-/ uncertain	[st̪]	/-ṣt̪-/ ≡ {/-t̪h-/ ^{alt} , (/d̪h-/) ^{alt} } The form /-t̪h-/ ^{alt} may have been used in the absence of a conjunct grapheme, although this is uncertain.
/-ṣt̪h -/	/-ṣt̪h-/ uncertain	[st̪ ^h]	/-ṣt̪h-/ ≡ {/-t̪h-/ ^{adv} (/th-/ , /st-/ , /t̪h-/) ^{tr} } It is uncertain whether the single retroflex stop records the retained or assimilated cluster. The allographs may record transitional pronunciations
/-st-/	/-st-/	[st]	/-st-/ ≡ {/-st-/ ^{adv} (/th-/) ^{alt} } This form is stable. The alternate form is possibly derived from a source text.
/-sth-/	/-tth-/ uncertain	[t̪ ^h] uncertain	/-tth-/ ≡ {/-t̪h-/ ^{adv} , /-t̪-/ ^{tr} , (/st-/ ^{tr} , /-th-/ ^{adv})} The advanced form /-th-/ ^{adv} is uncommon, while most reflexes appear to reflect alternate forms, recorded differently by various scribes.

The following sibilant + labial stop developments are all based on a small sample size.			
/-śp-/	/-śp-/	[śp]	/-śp-/ ≡ /-śp-/ ^h In this case the cluster is retained.
/-ṣp-/	/-śp-/	[śp]	/-śp-/ ≡ /-śp-/ ^{adv}
/-sp-/	/-śp-/	[śp]	/-śp-/ ≡ {/-śp-/ ^{adv} (/phr-/ , /-p-/) ^{alt} }

8.3.3.2 Fricative + nasal clusters

Examples: Fricative + nasal clusters occur both initially and medially in OIA. Gāndhārī reflexes are recorded for seven of the nine OIA clusters, and these are presented in Table 8.3.3.2a, along with a selection of typical examples from the source materials.

Table 8.3.3.2a: Gāndhārī reflexes of fricative + nasal clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-hm-	-(m)m-	Aś-H	<i>brāhmaṇa-</i> > <i>bramaṇa-</i>	3 BCE
		BLS1-2	<i>brāhmaṇa-</i> > <i>bramaṇa-</i>	1 CE
		RS-19	<i>brahma-</i> > <i>brama-</i>	2 CE
		RS-20	<i>brāhmaṇaḥ</i> > <i>bramaṇo</i>	
		InsK-M	<i>brāhmaṇa-</i> > <i>bramadata-</i>	2 CE
	-mh-	InsK-M	<i>brāhmaṇa-</i> > <i>bramhaṇa-</i>	2 CE
-śn-	-ś-	BLS4-1	<i>praśnena</i> > <i>pra[śe]ṇa</i>	1 CE
śm-	śp-	BLS1-1	<i>śmaśāna-</i> > <i>śpaśaṇa-</i>	1 CE
-śm-	-sp-	BLS2-2	<i>kaśmīra</i> > <i>[kaspīr.]</i>	1 CE
-ṣṇ-	-ṣ-	BLS2-1	<i>kṛṣṇaḥ</i> > <i>kriṣo</i>	1 CE
		BLS4-1	<i>akṛṣṇa-</i> > <i>akriṣa-</i>	
	-ṣ̄-	BLS2-1	<i>kṛṣṇam</i> > <i>kriṣo</i>	1 CE
		BLS4-1	<i>akṛṣṇa-</i> > <i>akriṣa-</i>	
		BLS9-1	<i>ūṣṇam</i> > <i>u[ṣ̄]a</i> , <i>trṣṇām</i> > <i>taṣa</i>	
		BCS5-1	<i>ūṣṇa-</i> > <i>uṣa-</i>	2 CE
	-ś-	InsK-M	<i>Kṛṣṇayasas-</i> > <i>Kriṣayaśa-</i>	1 CE
	-rṇ-	BLS1-2	<i>kṛṣṇaḥ</i> > <i>krirṇo</i>	1 CE

		BLS2-2	- <i>asmin</i> > - <i>ami</i>	
		InsK-K	- <i>asmi(n)</i> > - <i>ami</i>	2 CE

Other OIA combination not recorded in the Gāndhārī source materials are *hn* and *hṇ*.

Overview: Reflexes of clusters involving non-labial nasals are mostly represented as single sibilants, but often involving use of an overscore diacritic (§ 5.3.1). In word initial cases the typical development is *sn-* > *ṣ-*, while in medial positions the *-ṣṇ-* cluster mostly gives rise to reflexes of *-ṣ-* or *-ṣ̄-*, and the *-śn-* cluster to *-ṣ-* in a single example. There do not appear to be any conjunct *akṣaras* for S{n, ṇ} clusters in the Kharoṣṭhī script (§ 5.3.4)²⁶³, and understanding these reflexes depends on interpretation of the diacritic. Reflexes of clusters involving the labial nasal *m* are more common, and assimilate in two ways. The glottal fricative *h* is typically assimilated in *-hm-* > *-(m)m-*, while the most common development of the *Sm* clusters involves the loss of nasality and oral release of the occlusion in *-Sm-* > *-Sp-*, although assimilation as a geminate sibilant or geminate labial nasal are also seen in *-sm-* > *{-s(s)-, -(m)m-}*. Although characters for *Sm* clusters are recorded in other documents, no reflexes are recorded in the source materials, while *Sp* clusters are recorded using combined characters in the Kharoṣṭhī script (§ 5.3.4).

Discussion: The Gāndhārī reflexes of clusters involving non-labial nasals include word initial *sn-* > *ṣ-* and word medial *-ṣṇ-* > *{-ṣ-, -ś-, -ṣ̄-, -rṇ-}* and *-śn-* > *-ṣ-*. Clusters of the S{n, ṇ} type are retained in many of the modern Dardic languages (Baums 2009: 173)²⁶⁴, and it is suggested that the diacritic marked *Ṣ* reflexes indicate preserved clusters in Gāndhārī, used in the absence of conjunct *akṣaras* for the S{n, ṇ} clusters. An alternative interpretation of the superscript stroke is that this may indicate a geminate sibilant, as in *ss* or *ṣṣ* for example. This seems unlikely for three reasons, however: firstly, that geminates do not occur in word initial positions in OIA (§ 4.6), secondly, other geminate sibilant reflexes in Gāndhārī have been recorded using unmodified unweakened *akṣaras* (§§ 8.3.3.4.3, 8.3.3.4.4), and thirdly, in the absence of a conjunct *akṣara* for these clusters, the scribes would have developed an alternate diacritic marked form, all lending support to the interpretation of a retained OIA cluster. The suggested case of *sneha-* > *[si](^{*}ne)ho* appears to be an allophonic transcription of the retained cluster involving epenthetic *i* (§ 10.2.1), while it is also suggested that the *-ṣṇ-* > *-rṇ-* reflexes could indicate an aspirated

²⁶³ Although a character suggested as *ṣṇa*, in the hand of BL scribe 2, is recorded by Lenz (2010: 18), this is not included in lists of reflexes written by this scribe (Lenz 2003: 131; 2010: 28-29).

²⁶⁴ Refer to comments in footnote 259 and discussion in § 2.2.6 and § 8.2.6.3 regarding Dardic languages.

geminate nasal similar to the Pali reflexes (Baums 2009: 173). This is consistent with the use of preconsonantal *r* to mark geminates in other cases in the Kharoṣṭhī script (e.g., § 8.3.4.1.1). The unmarked single retroflex sibilant reflex *-ṣ-* of cluster *-ṣṇ-* is used alternately with the diacritic marked *-ṣ̄-* by two of the BL scribes, quite possibly due to scribal laxity, while *-ś-* is recorded in a single inscription only.

The Gāndhārī reflexes of clusters involving sibilants and the labial nasal *m* include word initial *śm-* > *sp-* and *sm-* > {*sp-*, *śp-*, (*sv-*)}, while medial reflexes include *-śm-* > *-sp-*, *-ṣm-* > {-*sp-*, *-śp-*, *-sp-*, *-ṣ-*} and *-sm-* > {-*sp-*, *-śp-*, *-s-*, *-(m)m-*, (*-sv-*)}. The predominant reflex involves loss of nasality to develop the labial stop *p* in a minor articulatory adjustment. The articulation of SN clusters involves a period of frication during the oral occlusion of the nasal, similar to that which occurs in the ST clusters (§ 8.3.3.1), and while they might be expected to possess a similar level of stability, the place and manner cues of nasals are less well-defined than for stops (§ 6.1), which may account for the loss of nasality in these reflexes. The main divergence to this pattern is regressive assimilation of the glottal fricative *h* by the labial nasal *m*, as for example in cases of Skt. *brahma-* > G *brama-*. This diverges from the typical Pali *-mh-* reflex involving metathesis of the fricative (Oberlies 2019: 149), as in Skt. *jihma-* > P *jimha-* for example (Suzuki 2002a: 106), although aspiration might also be interpreted in the Gāndhārī reflexes in the absence of a character for an aspirated labial nasal. A similar development also occurs in a small proportion of cases involving the dental sibilant in Skt. *-asmin* > G *-ami*.

It has been suggested that the variable development of *-ṣm-* > {-*s*, *ś*, *s*}*p-* in reflexes of Skt. *āyusman* by the RS scribe may derive from dialectal variants in the source texts, although different forms also occur within individual scrolls (Silverlock 2015: 267). The historical spelling is most probably *-sp-*, while *-śp-* appears to reflect a later palatalisation. Changes to the word initial vowels in developments such as Skt. *āyusmataḥ* > G *aśpado* may have led to phonetic uncertainty through the loss of the high vowel *u*, which triggers the OIA retroflexion of *s* in *āyusman*. Other reflexes such as *aīśpa* retain the high vowel *i* for *y* however, but this might also trigger palatalisation. A similar pattern is seen in the variable development of *-sm-* > {-*s*, *ś*}*p-* in the many cases involving variations on the Sanskrit pronominal locative ending *-asmin* and verbal form *asmi*. The palatal sibilant was used regularly by BL scribe 1²⁶⁵, while the RS scribe has mostly

²⁶⁵ Allon 2001: 95; Salomon 2008: 124-125.

retained the dental sibilant. The development of Skt. *āyasmā* > *aiśpa* in RS-19 appears to be an analogical spread from *āyusmā* > *aiśpa* in the same scroll, and the variation between these reflexes may be down to scribal habits as much as the source documents. The development of *-sm-* > *-s-* alternates with *-sp-* in similar words in the Aśokan inscriptions, also in the hands of BL scribe 9 and the RS scribe, suggesting a form of scribal shorthand rather than assimilation to geminate *-s(s)-*. The development of Skt. *asmi* > G *mi* is recorded by BL scribe 1 in the same texts as the *śpi* reflex, and may reflect copying of the common MIA development as *-mh-* from source texts, alternated with corrected Gāndhārī spellings in other places. Similar developments in some of the Kharoṣṭhī inscriptions may also reflect these patterns.

Summary: a summary of the Gāndhārī reflexes of OIA fricative + nasal clusters is presented in Table 8.3.3.2b. While the phonetic developments appear to be straightforward in most cases, the reflexes are represented in a variety of ways.

Table 8.3.3.2b: Summary of fricative + nasal cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
Word initial: the most common patterns are summarised as S_1n- > S_1n- and S_1m- > S_1p-			
/śn-/	/śn-/	[çn]	/śn-/ \equiv /ś-/ ^h The diacritic marked /ś-/ ^h has been interpreted as a representation of the preserved <i>Sn</i> clusters.
/sn-/	/sn-/	[sn]	/sn-/ \equiv {/ś-/ ^h , /sin-/ ^h }
/śm-/	/śp-/	[çp]	/śp-/ \equiv /śp-/ ^{adv}
/sm-/	/sp-/	[sp]	/sp-/ \equiv {/sp-/ ^{adv} , /śp-/ ^{alt} , (/sv-/ ^{alt})} The alternate forms result from palatalisation and lenition of the typical reflex.
Word medial: the homorganic coronal clusters appear to be retained: as in -SN- > -SN-. The labial nasal clusters develop as - S_1m- > - S_1p- , although the sibilants are interchanged orthographically.			

/-hm-/	/- mm-/	[mm]	/- mm-/ \equiv {/-m-/ ^{adv} , (/mh-/) ^{alt} } The alternate form appears to be a borrowing from another MIA dialect.
/-śn-/	/-śn-/	[çn]	/-śn-/ \equiv /-ś-/ ^h (from a single example only) The diacritic marked /ś-/ ^h has been interpreted as a representation of the preserved Sn clusters.
/-ṣṇ-/	/-ṣṇ-/	[ʃn]	/-ṣṇ-/ \equiv {/-ś-/ ^h , (/ṣ-/ ^{alt} , (/ś-/ ^{alt}), (/rṇ-/) ^{alt} } The alternate sibilant forms may be due to scribal laxity, while /-rṇ-/ ^{alt} may represent the MIA /-(n)nh-/ geminate reflex from a source text.
/-śm-/	/-śp-/	[çp]	/-śp-/ \equiv /-sp-/ ^{alt} (from a single example only) /-sp-/ ^{alt} and /-śp-/ ^{adv} are often interchanged as allographs in the reflexes of -Sm- clusters.
/-ṣm-/	/-ṣp-/	[ʃp]	/-ṣp-/ \equiv {/-ṣp-/ ^{adv} , /-śp-/ ^{alt} , /-sp-/ ^{alt} , (/ś-/ ^{alt})} /-sp-/ ^{alt} and /-śp-/ ^{alt} are often interchanged as allographs in the reflexes of -Sm- clusters.
/-sm-/	/-sp-/	[sp]	/-sp-/ \equiv {/-sp-/ ^{adv} , /-śp-/ ^{alt} , /-s-/ ^{alt} , /-m-/ ^{borr} , (/sv-/ ^{alt})} /-sp-/ ^{adv} and /-śp-/ ^{alt} are often interchanged as allographs in the reflexes of -Sm- clusters. The allograph /-s-/ is possibly a scribal shorthand, while /-m-/ appears to be a borrowed form and /-sv-/ records lenition.

8.3.3.3 Fricative + fricative clusters

Only three fricative-fricative clusters are recorded in the OIA sources (Table A3.1), which are geminates of the three sibilants, as in -śś-, -ṣṣ- and -ss-. These occur across syllable boundaries in the Ṛgveda, as in *rajas-su* and *haviṣ-ṣu* for example (Kobayashi 2004: 45-46), while -śś- appears to be rare. No reflexes of these clusters are recorded in the Gāndhārī source materials. Geminate sibilants do however develop in Gāndhārī as reflexes of sibilant plus semi-vowel clusters (§§ 8.3.3.4.3, 8.3.3.4.4).

8.3.3.4 Fricative + semi-vowel clusters

8.3.3.4.1 Fricative + *l* clusters

Examples: Gāndhārī reflexes of OIA fricative + *l* clusters recorded in the source materials, along with the recorded examples, are presented in Table 8.3.3.4.1a.

Table 8.3.3.4.1a: Gāndhārī reflexes of fricative + *l* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-hl-	-hal-	InsK-M	<i>Bāhlīka-</i> > <i>Bahalia-</i>	1 CE
śl-	śil-	BLS2-1	<i>ślokena</i> > <i>śilogano</i>	1 CE

One other OIA combination not recorded in the Gāndhārī source materials is *ṣl*.

Overview: In the two examples listed these clusters appear to develop through the insertion of an epenthetic vowel. As discussed in § 10.2.1 these often appear to represent allophonic pronunciation, and can be interpreted as representing retained clusters.

8.3.3.4.2 Fricative + *ṽ* clusters

Examples: All of the OIA fricatives combine with *ṽ* in both word initial and medial clusters. Gāndhārī reflexes of OIA fricative + *ṽ* clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.3.4.2a.

Table 8.3.3.4.2a: Gāndhārī reflexes of fricative + *ṽ* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-hv-	-(b)bh-	BLS1-2	<i>jihvā</i> > <i>jibha</i>	1 CE
		RS-20	<i>jihvā</i> > <i>-cibha-</i>	2 CE
śṽ-	śṽ-	InsK-M	<i>Śved(r)avarma</i> > <i>Śved(r)avarma</i>	2 CE
-śṽ-	-śp-	InsK-M	<i>viśva-</i> > <i>viśpa-</i>	1 BCE
			<i>Īśvaraka-</i> > <i>Iśparaka-</i>	1 CE
			<i>Viśvāmitra-</i> > <i>Viśpamitra-</i>	2 CE
			<i>Aśvayuj-</i> > <i>Aśpañu-</i>	3 CE
		BLS1-1	<i>aśvamūtrena</i> > <i>aśpamutrena</i>	1 CE
	-sp-	BLS4-1	<i>śāśvataḥ</i> > <i>saspado</i>	1 CE

sv-	sp-	Aś-H	svarga- > spagra-	3 BCE
		Aś-M	svāmi-ka > spamika	
		BLS4-1	svakayā > spayae	1 CE
		RS-12	svayam > [spa]ya	2 CE
	śp-	BLS2-1	svakam > śpagam	1 CE
		RS-5	svedaḥ > śpeda	2 CE
-sv-	-sv-	BCS5-1	svayam > śpa[ho], svabhāvatā > śpabhavasa	
		InsK-M	-svāmī- > -svami-	2 CE
			-svāmī- > -svami-	3 CE
	-śp-	BLS1-1	visvaram > viśparo	1 CE
		RS-5	parisvedāni > pariśpeidaṇa	2 CE
	-sp-	BCS5-2	asvāmika- > aspamia-	

The other OIA combination not recorded in the Gāndhārī source materials is śv.

Overview: The majority of recorded reflexes involve the OIA śv and sv clusters, which typically develop as Sp in both word initial and medial positions. There is some alternation of ś for s in the dental cluster reflexes, while clusters are preserved in a small number of cases. The glottal fricative -hv- cluster appears to develop as a geminate stop -(b)bh- based on two available examples. A version of sva- (𑀭𑀸𑀓) used by the BC scribe 5 has a similar diacritic flourish to that used for tva- (§ 5.3.2.1), but its use appears limited to reflexes of Skt. *tasmāt* > G *tasva* (Schlosser 2016: 83). This suggests that sv clusters were not retained as they could otherwise have been represented in this way, while the Sp clusters were commonly represented as a conjunct character (§ 5.3.4). Three examples of retained Sv clusters are listed from the Kharoṣṭhī inscriptions, but these involve a proper noun and the common term -svami-. It is difficult to discern the ligatures from the available images²⁶⁶.

Discussion: The Gāndhārī Sv > Sp reflexes develop through occlusion of the voiced labial-dental approximant v as the voiceless labial stop p, although it has been suggested, as discussed below, that this may have developed through an intermediate stage of Sm. The Sanskrit semi-vowel v appears to have been pronounced as a bilabial in the early

²⁶⁶ Konow 1929: 127, Plate XXIII; 152, Plate XXVIII; 173, Plate XXXV.

stages of the language, possibly similar to the voiced labial-velar fricative [w]²⁶⁷, but appears to have moved closer the labio-dental approximant [v], at least in some dialects, in later periods. This is akin to the Hindi labiodental approximant, which is pronounced with a more open and rounded lip pattern (Ladefoged and Maddieson 1996: 323-326). It has been suggested that a more precise articulatory effort is required to maintain the correct closure for frication than for a stop (Kirchner 1998, 111), so that an occlusive development such as Sv > Sp may lessen the articulatory effort, especially if an element of frication was retained in the pronunciation of v at that time. It may be easier to close the lips fully, rather than adjust to a controlled fricative closure in the lips following the fricative sibilant, and such a change may be viewed as one of lenition rather than fortition.

Although represented in a limited sample, the medial -Sp- clusters are typically preserved (§ 8.3.3.1), while the initial and medial Sm clusters also develop as Sp (§ 8.3.3.2), so that the Sp, Sm and Sv clusters all merge as -Sp- in medial positions in Gāndhārī, while both Sm and Sv merge as Sp- in word initial positions. In assessing these developments Baums (2009: 174-177) has suggested that an intermediate stage merger occurred between Sv and Sm prior to the final merger with Sp, which may have developed as either Sv > Sm > Sp, or Sm > Sv > Sp. There is, however, little direct evidence for these intermediate stages compared to the significant record of direct developments. Two atypical cases of sm > sv, one initial and one medial, are recorded in Table 8.3.3.2a, while Sv > Sm is not recorded at all in Table 8.3.3.4.2a. Developments of sm > sv are also recorded in Central Asian documents²⁶⁸, but these also appear to be atypical. And while the development of -tm- > -tv- among the stop clusters (§ 8.3.1.2) may lend support the idea of a Sm > Sv > Sp staged development, this involves lenition of the nasal occlusion in release from the closure of the initial stop. On the other hand, Sv > Sp develops following the partial closure of the initial frication, while Sm > Sp also involves a simple loss of nasality. At the same time there is significant intermixing between the nasal consonants p, m and v in Gāndhārī, and although not supported clearly by the available record, it is possible that the diachronic developments included intermediate stages of either form.

While the sibilant characters are often intermixed by the scribes in these reflexes, the three sibilant sounds were in fact retained in Gāndhārī (§ 8.2.3). Orthographic mixing

²⁶⁷ Macdonell 1916: 17; Allen 1953: 57.

²⁶⁸ Burrow 1937: 20; Brough 1962: 102-103.

of the sibilants appears to reflect scribal preferences as well as source text divergences. Development of the glottal fricative cluster *-hv-* > *-(b)bh-* reflects a similar development as the voiced stop *b* and subsequent metathesis of the glottal fricative as aspiration. The retention of *śv-* > *śv-* seen in the Kharoṣṭhī inscriptions appears to reflect the historic spelling of a proper noun, while *-sv-* > *-sv-* suggests retention of the historic spelling for the common word *svāmin*.

Summary: a summary of the Gāndhārī reflexes of OIA fricative + *v* clusters is presented in Table 8.3.2.4.2b.

Table 8.3.3.4.2b: Summary of fricative + *v* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
Word initial: the most common pattern is summarised as S_1v- > S_1p- .			
/śv-/	/śv-/	[çv]	/śv-/ \equiv /śv-/ ^h This case is based on a single example involving a proper noun
/sv-/	/sp-/	[sp]	/sp-/ \equiv {/sp-/ ^{adv} , /śp-/ ^{alt} } /sp-/ and /śp-/ are interchanged as allographs in the reflexes of -Sv- clusters.
Word medial: in the majority of cases clusters develop as S_1v > S_1p , although there are variations in the recording of the sibilant and the clusters are preserved in some cases.			
/-hv-/	/-(b)bh-/	[bb ^h]	/-(b)bh-/ \equiv /-bh-/ ^{adv} (based on two examples only)
/-śv-/	/-śp-/	[çp]	/-śp-/ \equiv {/-śp-/ ^{adv} , (/sp-/ ^{alt})}
/-sv-/	/-sp-/	[sp]	/-sp-/ \equiv {/-śp-/ ^{alt} , /-sv-/ ^h } The historic allograph was used in a common word, while /sp-/ and /śp-/ are interchangeable allographs in the reflexes of -Sv- clusters.

8.3.3.4.3 Fricative + y clusters

Examples: All OIA fricatives combine with y in both word initial and medial clusters (Table A3.1). Gāndhārī reflexes of OIA fricative + y clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.3.4.3a.

Table 8.3.3.4.3a: Gāndhārī reflexes of fricative + y clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-hy-	-hiy-	InsK-M	<i>mahyam</i> > <i>mahiya</i>	2 CE
	-h(h)-	BLS1-1	* <i>grhya</i> > <i>ghriha</i>	
	-ś-	BLS1-3	<i>vigāhya</i> > <i>vighaśa</i> , <i>āruhya</i> > <i>aruśu</i>	1 CE
	-ś- -j-	BLS4-1	<i>dahyamānam</i> > <i>daśamaṇa</i> , <i>dajamaṇa</i>	
	-z-	RS-5	<i>guhya</i> > <i>guza</i> (secondary voicing)	2 CE
	(y)y	RS-19	<i>-uhyate</i> > <i>avuyadi</i>	2 CE
-śy-	-śiy-	Aś-H	<i>prativeśya</i> > <i>prativeśiya</i> -	3 BCE
	-ś(ś) -	InsK-M	<i>Kāśyapīya</i> > <i>Kaśavia</i> -	1 CE
		BLS1-1	<i>praṇaśyati</i> > <i>praṇaśadi</i>	1 CE
		BLS4-1	<i>paśyati</i> > <i>paśadi</i>	
		RS-5	<i>paśyati</i> > <i>paśaḍi</i>	2 CE
		RS-19	<i>paśyati</i> > <i>paśaso</i>	
		BCS5-1	<i>avaśyam</i> > <i>avaśa</i>	
		InsK-M	<i>Kāśyapīya</i> > <i>Kaśaviya</i> -	3 CE
	-ṣ-	BLS2-1	BHS <i>karāpiśye</i> > <i>karavi[ṣa]</i>	1 CE
	-ṣ-	BCS5-1	<i>deśyamāna</i> > <i>deṣamana</i> -	2 CE
-ṣy-	-ṣy-	InsK-M	<i>-śiśya</i> > <i>-śiśya</i> -	2 CE
	-ś(ś)-	Aś-H	<i>-iśyati</i> > <i>-iśati</i>	3 BCE
		Aś-M	<i>manuṣya</i> > <i>manuśa</i> -	
		BLS1-1	<i>kariṣye</i> > <i>kariśe</i> and many similar,	1 CE
		BLS2-2	<i>manuṣya</i> > <i>manuśa</i> , <i>utthāpayiṣyati</i> > <i>uthavi[ṣa]</i>	
		BLS4-1	<i>duṣyati</i> > <i>duśadi</i>	
		BLS9-1	<i>manuṣya</i> > <i>manuśa</i>	
		RS-5	<i>mariṣyāmi</i> > <i>[m](*)riśami</i>	2 CE

number of clusters remain stable or are resolved through the insertion of epenthetic *i*, while a small number of variant reflexes such as \underline{s} and \bar{s} are also recorded. A number of reflexes are recorded for the OIA medial *-hy-* cluster, which include assimilation of *y*, insertion of epenthetic *i*, and the use of different palatal consonant symbols. While there do not appear to be any examples of retained clusters in the manuscript sources, conjunct ligatures are used to represent the *-sya-* and *-ṣya-* in the Kharoṣṭhī inscriptions. These involve a leftward sweeping diacritic mark attached to the base of the sibilant *akṣara*, as in (𑀓𑀲) and (𑀓𑀲𑀭) respectively for example, from the second century CE Sui Vihār copper plate inscription²⁶⁹. Although curving slightly differently, these downward strokes are similar to the markings used to represent the *Ty* clusters in other inscriptions (§ 5.3.2.1).

Discussion: The assimilation of OIA medial *Sy* clusters as geminate sibilants appears to be the normal development in Gāndhārī²⁷⁰, similar to the typical MIA reflexes (§ 9.1.2). This is consistent with the similar resolution of *Ty* and *Ny* clusters as the *TT* and *NN* geminates (§§ 8.3.1.4.3; 8.3.2.4.3), and while the *Sr* clusters also assimilate as geminate sibilants (§ 8.3.3.4.4), the other *CR* clusters are either retained or assimilated in different ways. Possible reasons for the relative weakness of the *y* semi-vowel have been discussed previously and are explored further in § 9.3, and it is possible that these reflexes represent an early spread of the wider MIA assimilation patterns among the weaker semi-vowels in Gāndhārī. This suggests that both *y* and *r* are the most open and least perceptually reliable of the semi-vowels when following a sibilant.

While the occasional recording of preserved clusters may be interpreted as historic spellings, the common insertion of epenthetic vowels in the word initial *sy-* clusters is possibly an allophonic recording of the preserved cluster (§ 10.2.1). Since geminates do not occur word initially, the diacritic \underline{s} - reflex could also represent the retained cluster, at least in this situation. Similar interpretations might be applied to the medial palatal cluster reflexes of *-śy-* > *-śiy-* / *-ṣ-* recorded in the Aśokan inscriptions and by the BC scribe 5 respectively, but these are isolated cases and the recording of single *-ś-* is the most common reflex in this case. A similar pattern is also seen in the medial dental sibilant cluster reflexes of *-sy-* > *-s-*, which mostly occurs in reflexes of the OIA genitive endings *-asya*, although underscore dental sibilant reflex *-ṣ-* is sometimes used by the RS scribe.

²⁶⁹ Copied from Konow 1929: Plate XXVI, following p.140.

²⁷⁰ Allon 2001: 96-97; Salomon 2008: 125; Baums 2009: 178; Silverlock 2015: 269.

While this reflex might also be interpreted as indicating the retained cluster, similar to its word initial use by the same scribe in *syāt* > *śa*, it has been suggested this may indicate a further phonetic development in which the sibilant was shortened and voiced, with compensatory lengthening of the preceding vowel, as in *-aśa* = *-āza* (Baums 2009: 178). Both reflexes have been used by the RS scribe in the same documents, however, without a clear distinction between alternate uses, suggesting that this could also have been used as an alternate graphic marking of the geminate.

Developments of the OIA *-hy-* cluster indicate retention in the epenthetic *-hiy-* reflex, assimilation as *-h(h)-* or *-(y)y-*, or palatalisation as *-ś-*, *-z-* or *-j-*, suggesting an unclear phonetic realisation. Both *h* and *y* are relatively weak in the scale of consonant strengths making them more prone to lenition (Gordon 2016: 153), and the loss of either consonant is understandable in this context. Realisation as some form of palatal fricative also seems possible (Campbell 2013: 35-36).

Summary: a summary of the Gāndhārī reflexes of OIA fricative + *y* clusters is presented in Table 8.3.3.4.3b.

Table 8.3.3.4.3b: Summary of fricative + *y* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/sy-/	/sy-/	[sj]	/sy-/ ≡ {/siy-/ ^h , (/ś-/ ^{alt} , /s-/ ^{alt})} The epenthetic form is an allophone of the retained historic form, while the <i>ś-</i> alternate may be a way of representing the retained cluster.
Word medial: in the majority of cases clusters develop as <i>S₁y</i> > <i>S₁S₁</i> , although the retroflex sibilant is typically palatalised.			
/-hy-/	/-hy-/ uncertain	unclear palatalised	/-hy-/ ≡ {/-hiy-/ ^h , (/ś-/ ^{alt} , /z-/ ^{alt} , /j-/ ^{alt} , /h-/ ^{alt} , /y-/ ^{alt})} The development of this cluster is unclear. While the historic form may have been retained, the alternate reflexes indicate some form of palatalisation.

/-śy-/	/-śś-/	[çç]	<p>/-śś-/ ≡ {/-ś-/^{adv}, (/ś-/^{alt}, /-śiy-/^h)}</p> <p>The alternate form is assumed to represent the geminate sibilant, while the epenthetic variation is assumed to be an allophonic transcription of the retained historic form.</p>
/-ṣy-/	/-śś-/	[çç]	<p>/-śś-/ ≡ {/-ś-/^{adv}, (/ś-/^{alt}, /-(ṣ)ś-/^{alt}, /-ṣy-/^h)}</p> <p>The geminate sibilant is palatalised in the common reflex, while the retroflex sibilant is retained in one case. The historic form is found in one inscription, while the diacritic may reflect gemination or a phonetic variation.</p>
/-sy-/	/-ss-/	[ss]	<p>/-ss-/ ≡ {/-s-/^{adv}, /-ś-/^{alt}, (/sy-/^h, /-siy-/^h)}</p> <p>The alternate reflex may represent an advanced phonetic form as a voiced single sibilant, or otherwise mark the gemination.</p>

8.3.3.4.4 Fricative + *r* clusters

Examples: All of the OIA fricatives combine with *r* in word initial and medial clusters, apart from retroflex ṣ which does not occur word initially and is infrequent in medial positions. Gāndhārī reflexes of OIA fricative + *r* clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.3.4.4a.

Table 8.3.3.4.4a: Gāndhārī reflexes of fricative + *r* clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>hr-</i>	<i>r-</i>	BLS1-2	<i>hrada-</i> > <i>rada-</i>	1 CE
<i>śr-</i>	<i>śr-</i>	Aś-H	<i>śrāvaka-</i> > <i>śravaka-</i> , <i>śramana-</i> > <i>śramana-</i>	3 BCE
		Aś-M	<i>śruṇu-</i> > <i>śruṇa-</i>	3 BCE
		InsK-M	<i>śramaṇa-</i> > <i>śramaṇa-</i>	1 BCE
			<i>śrāvaṇa-</i> > <i>śravaṇa-</i>	1 CE
		BLS4-1	<i>śreṣṭha-</i> > <i>śreṭh[a]-</i>	1 CE
		RS-19	<i>śrutvā</i> > <i>śrutva</i>	2 CE
	<i>rś-</i>	BLS1-2	<i>śruta-</i> > <i>rśoda-</i>	1 CE

<i>-śr-</i>	<i>sr-</i>	BLS1-1	<i>śreṣṭhin-</i> > <i>sreṭhi-</i>	1 CE
	<i>śir-</i>	BLS1-1	<i>śrīmataḥ</i> > <i>śirimad[u]</i>	1 CE
	<i>ś-</i>	Aś-H	<i>śramaṇa</i> > <i>śamana</i>	3 BCE
		BLS1-2	<i>Śrāvastyām</i> > <i>Śavastie</i>	1 CE
		RS-5	<i>Śrāvastī</i> > <i>Śava[st]i</i>	2 CE
		RS-12	<i>śrutvā</i> > <i>śutva</i>	
		RS-19	<i>śrutvā</i> > <i>śutva</i>	
		RS-20	<i>Śrāvastī-</i> > <i>Śavasti-</i>	
	<i>ṣ-</i>	InsK-M	<i>śraddhā-</i> > <i>[ṣa]dha-</i> , <i>śruta-</i> > <i>[ṣu]ta-</i>	1 CE
			<i>śramaṇa-</i> > <i>-ṣamaṇa-</i> , <i>śrāvaka-</i> > <i>ṣavaa-</i>	2 CE
			<i>śrāvaka-</i> > <i>ṣavaa-</i>	3 CE
		BLS1-1	<i>śramaṇam</i> > <i>ṣamaṇo</i>	1 CE
		BLS1-2	<i>śrutvā</i> > <i>ṣutvaṇa</i>	
		BLS2-1	<i>śramaṇena</i> > <i>ṣamanano</i> , <i>śrūyate</i> > <i>ṣuyadi</i>	
		BLS2-2	<i>śramaṇa-</i> > <i>ṣamaṇa-</i> , <i>śrūyate</i> > <i>ṣuyadi</i>	
		BLS4-1	<i>śrāvakāḥ</i> > <i>ṣavaga</i> , <i>śrutvā</i> > <i>ṣutva</i>	
		RS-5	<i>śraddhasya</i> > <i>ṣadhasa</i>	2 CE
		RS-19	<i>śraddhayā</i> > <i>ṣadha</i>	
		RS-20	<i>śramanāḥ</i> > <i>ṣamaṇa</i>	
	<i>s-</i>	BLS1-2	<i>śrotra-</i> > <i>suda-</i>	1 CE
		RS-20	<i>śrotra-</i> > <i>sotro-</i>	2 CE
	<i>-śr-</i>	Aś-H	<i>śuśrūṣā</i> > <i>suśruṣa</i>	3 BCE
		BLS2-1	<i>*pariśravam</i> > <i>pariśravo</i>	1 CE
		RS-12	<i>aśrauṣīt</i> > <i>[a]śroṣi</i>	2 CE
	<i>-rś-</i>	BLS1-1	<i>āśramaḥ</i> > <i>arśamu</i>	1 CE
		BLS1-2	<i>pratyaśrauṣuḥ</i> > <i>pracarśoṣu</i>	
	<i>-ś(ś)-</i>	Aś-H	<i>niśritam</i> > <i>niśite</i>	3 BCE
		RS-5	<i>aśru-</i> > <i>aśu-</i>	2 CE
		BCS5-1	<i>vyatimiśra-</i> > <i>vidimiśa-</i>	
	<i>-ś̄-</i>	BLS2-2	<i>pariśravaḥ</i> > <i>pariśave</i>	1 CE
	<i>-ṣ(ṣ)-</i>	BLS1-2	BHS <i>niśrita-</i> > <i>ṇiṣida-</i>	1 CE
		BLS9-1	<i>pariśrayāni</i> > <i>pariṣeāṇi</i>	
		RS-12	BHS <i>anuśrāvayensuḥ</i> > <i>aṇuṣavisu</i>	2 CE
		BCS5-1	<i>miśra-</i> > <i>miṣo-</i>	2 CE

	-ṣ-	BCS5-1	<i>miśra-</i> > <i>miṣo-</i>	2 CE
	-s-	Aś-M	<i>pariśrava-</i> > <i>-parisava-</i> (M)	3 BCE
			<i>miśra-</i> > <i>misa-</i>	
-sr-	-sr-	Aś-H	<i>parisrava</i> > <i>parisrave-</i> , <i>sahasra-</i> > <i>sahasra-</i>	3 BCE
		InsK-M	<i>sahasra-</i> > <i>sa[ha]s[ra]-</i>	1 CE
		RS-12	<i>āsravāḥ</i> > <i>asrava</i>	2 CE
	-s(s)-	Aś-H	<i>sahasra-</i> > <i>sahasa-</i>	3 BCE
		BLS1-2	<i>sahasra-</i> > <i>sahasa-</i>	1 CE
	-ṣ-	BLS4-1	<i>anāsravaḥ</i> > <i>aṇasava</i>	
		RS-5	<i>āsravānām</i> > <i>asavaṇa</i>	2 CE
		BLS4-1	<i>anavasrutaḥ</i> > <i>aṇavaṣuda</i>	1 CE
		RS-5	<i>āsravānām</i> > <i>[a]ṣavaṇa</i>	2 CE
		RS-19	<i>āsravebhyaḥ</i> > <i>aṣavehi</i>	

One other OIA combination not recorded in the Gāndhārī source materials is *ṣr*.

Overview: The development of these clusters is quite variable, seen in *śr*- > {*śr*-, *ś*-, *ṣ*-, (*śir*-, *rś*-, *sr*-, *s*-)} for the word initial clusters, and *-śr*- > {-*śr*-, *-rś*-, *-ś*-, *-ṣ*-, (*-ṣ̣*-, *-s*-, *-ṣ*-)} and *-sr*- > {-*sr*-, *-s*-, (*-ṣ*-, *-ṣ*-)} for the medial clusters, while *hr*- > *r*- is recorded in one example only. While a significant number of reflexes indicate retention of the clusters in both initial and medial positions, there are also a considerable number which suggest deletion of the *r* and reduction to a single sibilant, both initially and medially, with possible gemination in medial positions. Diacritic marked reflexes appear to be used on an occasional basis only, and may reflect transitional or allophonic forms. Post-consonantal *r* is typically recorded with a diacritic stroke to the base of the akṣara in the Kharoṣṭhī script, as in *śra*- (𑖦𑖻) and *sre* (𑖦𑖿) for example (§ 5.3.2.1).

Discussion: Unlike the *Tr* and *Nr* clusters which are mostly retained in Gāndhārī (§§ 8.3.1.4.4; 8.3.2.4.4), the *Sr* clusters are commonly reduced to single sibilants initially, and assimilated as geminates medially, represented by a single sibilant, while also being retained in many cases. This partial assimilation of the *Sr* clusters, along with the widespread assimilation of the *Sy* clusters (§ 8.3.3.4.3), suggests that *y* and *r* have the least reliable perceptual cues of the semi-vowels when following a sibilant. The reduction of word initial *śr*- > *ś*- reflects lenition of *r* in the prevocalic position and may have followed

its weakening from a tap to a continuant. The development of word medial *-Sr-* > *-S-* may also reflect a process of lenition, while the development of a geminate *-SS-* reflex might be explained in terms of progressive manner assimilation of the rhotic by the sibilant, similar to the MIA developments (§ 9.1.2). This has been explained in terms of the diachronic development of anticipatory consonant ‘doubling’ across the syllable boundary recorded in the Sanskrit gemination accounts, which is discussed further in § 9.3. This might develop as *-.Sr- > s.Sr- > -s.Sʳ- > S.S* in this case.

The original OIA *Sr* clusters are also retained in many examples, ranging from the third century BCE Aśokan inscriptions through to the BL and RS manuscripts in the first and second centuries CE, where some scribes have also used both advanced and retained forms. Since these clusters were generally assimilated to geminate sibilants in MIA and Pali, it is also possible that advanced forms are borrowings from another MIA dialect, while the historic forms represent the actual pronunciation, and in this situation the scribes may have been in two minds whether to write historic or borrowed forms. Their continued use over several centuries may also reflect spelling alternations which developed during the early stages of this sound transition and subsequently retained over several centuries. These variable records offer valuable insights into the processes of language change in progress.

The alternative development of *-śr- > -ṣ(ṣ)-*, recorded by both the BL and RS scribes, may be conditioned by the alveolar place of *r*, while the more consistent development of *-sr- > -s(s)-* suggests that an alveolar *r* has less power to retract the preceding dental *s* than as a true retroflex *r* (Baums 2009: 179).

Summary: a summary of the Gāndhārī reflexes of OIA fricative + *r* clusters is presented in Table 8.3.3.4.4b.

Table 8.3.3.4.4b: Summary of fricative + *r* cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/hr-/	/r-/	[r]	/r-/ \equiv /r-/ ^{adv} This case is based on a single example.
/śr-/	/ś-/	[ç]	/ś-/ \equiv {/ś-/ ^{adv} , /ś-/ ^{alt} , /śr-/ ^h } The historic form /śr-/ ^h may reflect historic alternations or borrowings, while occasional variants such as /rś-/ and /śir-/ also occur.
Word medial: the most common development appears to be -Sr- > -SS- although there are variations in the recording of the sibilant and the clusters are preserved in some cases. This appears to be a similar although partially developed variant of the typical MIA reflex where <i>r</i> is assimilated into a geminate sibilant.			
/-śr-/	/-śś-/	[ç]	/-śś-/ \equiv {/-ś-/ ^{adv} , /-ś-/ ^{alt} , /-śr-/ ^h , (/ś-/ ^{tr} , /-ś-/ ^{alt}) ^{alt} } The retroflex sibilant /-ś-/ ^{alt} appears to reflect conditioning by alveolar <i>r</i> , while /-śr-/ ^h may reflect historic alternations or borrowings. Occasional alternate forms such as /-ś-/ ^{tr} and /-ś-/ ^{alt} also occur.
/-sr-/	/-ss-/	[s]	/-ss-/ \equiv {/-s-/ ^{adv} , /-sr-/ ^h , (/ś-/ ^{tr} , /-ś-/ ^{alt}) ^{alt} } The historic form /-sr-/ ^h may reflect historic alternations or borrowings, while occasional alternate forms such as /-ś-/ and /-ś-/ ^{alt} also occur.

8.3.4 Semi-vowel plus clusters

The majority of OIA semi-vowel plus clusters are of the type *-r{T, N, S, R}-* in which *r* is the initial member preceding consonants having all other manners of articulation (Tables A3.1, A4.1). These are followed in frequency by *-lC-* clusters which mostly occur in combination with the labial stops and other semi-vowels, then *-vC-* clusters which mostly occur in combination with other semi-vowels, and finally *-yC-* clusters which are far less common. These clusters only occur medially except for *vy-* and *vr-*.

8.3.4.1 *l* + C clusters

8.3.4.1.1 *l* + stop clusters

Examples: Gāndhārī reflexes of OIA *l* + stop clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.4.1.1a.

Table 8.3.4.1.1a: Gāndhārī reflexes of *l* + stop clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>-lp-</i>	<i>-(p)p-</i>	Aś-H	<i>kalpa-</i> > <i>kapa-</i> , * <i>alpa-</i> > <i>apa-</i>	3 BCE
		BLS1-1	<i>kalpa-</i> > <i>kapa-</i>	1 CE
		BLS9-1	<i>-kalpaḥ</i> > <i>-gapo</i>	
		BCS5-2	<i>parikalpa-</i> > <i>pariapa-</i>	2 CE
		RS-12	<i>kalpāmi</i> > <i>kapemi</i>	2 CE
		RS-19	<i>saṃkalpa-</i> > <i>sakapa-</i>	
	<i>-rp-</i>	BLS2-2	<i>saṃkalpa-</i> > <i>sakarpa-</i> , uncertain	1 CE
		BCS5-1	<i>kalpa-</i> > <i>karpa-</i>	2 CE
<i>-lbh-</i>	<i>-(b)bh-</i>	BLS4-1	<i>apragalbhah</i> > (*a)[p](*)ag[a]bho	1 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *lk*, *lg*, *ld*, *lph* and *lb*.

Overview: Although the sample is mostly limited to reflexes of Skt. *-kalpa-* > G *-kapa-*, in the majority of cases the semi-vowel *l* is assimilated by the following labial stop *p*. This is typically recorded as a single unweakened stop assumed to represent a geminate, while the

development of the OIA *lbh* cluster involves a similar regressive assimilation of the semi-vowel to form a geminate stop. Combined character forms representing pre-consonantal *l* have been recorded on a coin inscription and in the Niya documents (Glass 2000: 132), where the character *lpa* (𐤙) is a combination of the forms used for *l* (𐤙) and *p* (𐤕), although no similar characters appear to be recorded in the inscriptional and manuscript sources published to date²⁷¹. The reflex of *-lp-* > *-rp-*, represented as (𐤕) (§ 5.3.2.2), has been recorded by BL scribe 2 and BC scribe 5.

Discussion: The *l* + labial stop clusters form geminate stops through regressive place and manner assimilation of the alveolar lateral approximant to form a geminate labial stop. The manner assimilation is consistent with the typical MIA strength-based assimilation (§ 9.2.2), and the place assimilation with the stronger perceptual cues and articulatory movements of the labial stop (§ 6.1). It has been suggested that the pre-consonantal *r* diacritic in the *rp* reflexes may have been used to mark the geminate stop reflex²⁷², although it is possible this was intended to mark a retained *l* (Schlosser 2016: 86). This may reflect the phonetic similarities between the two glides (§ 8.3.1.4.1), but represented using the simpler and more common *r* diacritic, or that the *r* comes from a northwestern "*r*-dialect". Although based on limited data and by no means certain, this raises the possibility of some level of retention of the cluster, similar to the conservative forms recorded in the Niya documents (Burrow 1937: 15, § 40).

8.3.4.1.2 *l* + nasal clusters

There is only one *l* + nasal cluster in OIA, which is recorded once as *-lm-* > *-(m)m-* in Skt. *kulmāṣaḥ* > G *k[u]ma[ṣa]* (RS-5, 2 CE) in the Gāndhārī source documents. While this single example does not offer any clear indication of the Gāndhārī developments, it is consistent with the relative perceptual and articulatory strengths of the labial nasal which appear favourable to regressive assimilation of the lateral approximant (§ 6.1). The *lm* cluster is recorded and written as a combined character in the Niya documents, as in (𐤙) for example (Glass 2000: 132), where it appears to have been retained as an archaic form (Burrow 1937: 15, § 40).

²⁷¹ As noted in § 5.3.4, a similar form of combined character for *lpa* has been observed in preliminary studies of the BL fragment 4 by Tim Lenz. This appears in G *alpamasa-* (T.Lenz, personal communication, September 11 and 18, 2023).

²⁷² Salomon 2000: 77; Lenz 2010: 33; Schlosser 2016: 86.

8.3.4.1.3 *l* + fricative clusters

There are two fricative clusters, *lś* and *lh* in OIA, neither of which is recorded in the Gāndhārī source materials.

8.3.4.1.4 *l* + semi-vowel clusters

Examples: There are three OIA *l* + semi-vowel clusters, and only reflexes of *-ly-* are recorded in the Gāndhārī source materials. These are presented in Table 8.3.4.1.4a along with a selection of typical examples.

Table 8.3.4.1.4a: Gāndhārī reflexes of *l* + semi-vowel clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>-ly-</i>	<i>-ly-</i> <i>-l(l)-</i>	InsK-M	* <i>Kamagulya-</i> > <i>Kamagulya-</i>	2 CE
		Aś-M	<i>kalyāṇa-</i> > <i>kalāṇa-</i>	3 BCE
		BLS1-1	<i>mālyam</i> > <i>malu</i> / <i>ma[l](*)o</i>	1 CE
		BLS4-1	<i>kauśalya-</i> > <i>kośala-</i> , * <i>kālyāṇa-</i> > <i>kalāṇa-</i>	
		BLS9-1	<i>kalyāṇa-</i> > <i>kalāṇa-</i>	
		BCS5-2	<i>śalya-</i> > <i>śala-</i>	2 CE

Other OIA combinations not recorded in the Gāndhārī source documents are *lv* and *ll*.

Overview: Although the sample is limited, the glide *y* is typically assimilated by the preceding liquid *l*, while the cluster appears to be retained in its conservative form in the proper noun from the second century CE Wardak vase inscription (Hultsch 1925: 170). The retention of post-consonantal *y* appears to be a common pattern in the Wardak vase inscription (§ 8.3.1.4.3), and *ly* is recorded using a similar wavy diacritic attached at the base of the preceding akṣara, as in (𑀭𑀮𑀹𑀲𑀳𑀴𑀵𑀶𑀷𑀸𑀹𑀺𑀻𑀼𑀽𑀾𑀿𑁀𑁁𑁂𑁃𑁄𑁅𑁆𑁇𑁈𑁉𑁊𑁋𑁌𑁍𑁎𑁏𑁐𑁑𑁒𑁓𑁔𑁕𑁖𑁗𑁘𑁙𑁚𑁛𑁜𑁝𑁞𑁟𑁠𑁡𑁢𑁣𑁤𑁥𑁦𑁧𑁨𑁩𑁪𑁫𑁬𑁭𑁮𑁯𑁰𑁱𑁲𑁳𑁴𑁵𑁶𑁷𑁸𑁹𑁺𑁻𑁼𑁽𑁾𑁿𑂀𑂁𑂂𑂃𑂄𑂅𑂆𑂇𑂈𑂉𑂊𑂋𑂌𑂍𑂎𑂏𑂐𑂑𑂒𑂓𑂔𑂕𑂖𑂗𑂘𑂙𑂚𑂛𑂜𑂝𑂞𑂟𑂠𑂡𑂢𑂣𑂤𑂥𑂦𑂧𑂨𑂩𑂪𑂫𑂬𑂭𑂮𑂯𑂰𑂱𑂲𑂳𑂴𑂵𑂶𑂷𑂸𑂺𑂹𑂻𑂼𑂽𑂾𑂿𑃀𑃁𑃂𑃃𑃄𑃅𑃆𑃇𑃈𑃉𑃊𑃋𑃌𑃍𑃎𑃏𑃐𑃑𑃒𑃓𑃔𑃕𑃖𑃗𑃘𑃙𑃚𑃛𑃜𑃝𑃞𑃟𑃠𑃡𑃢𑃣𑃤𑃥𑃦𑃧𑃨𑃩𑃪𑃫𑃬𑃭𑃮𑃯𑃰𑃱𑃲𑃳𑃴𑃵𑃶𑃷𑃸𑃹𑃺𑃻𑃼𑃽𑃾𑃿𑄀𑄁𑄂𑄃𑄄𑄅𑄆𑄇𑄈𑄉𑄊𑄋𑄌𑄍𑄎𑄏𑄐𑄑𑄒𑄓𑄔𑄕𑄖𑄗𑄘𑄙𑄚𑄛𑄜𑄝𑄞𑄟𑄠𑄡𑄢𑄣𑄤𑄥𑄦𑄧𑄨𑄩𑄪𑄫𑄬𑄭𑄮𑄯𑄰𑄱𑄲𑄳𑄴𑄵𑄶𑄷𑄸𑄹𑄺𑄻𑄼𑄽𑄾𑄿𑅀𑅁𑅂𑅃𑅄𑅅𑅆𑅇𑅈𑅉𑅊𑅋𑅌𑅍𑅎𑅏𑅐𑅑𑅒𑅓𑅔𑅕𑅖𑅗𑅘𑅙𑅚𑅛𑅜𑅝𑅞𑅟𑅠𑅡𑅢𑅣𑅤𑅥𑅦𑅧𑅨𑅩𑅪𑅫𑅬𑅭𑅮𑅯𑅰𑅱𑅲𑅳𑅴𑅵𑅶𑅷𑅸𑅹𑅺𑅻𑅼𑅽𑅾𑅿𑆀𑆁𑆂𑆃𑆄𑆅𑆆𑆇𑆈𑆉𑆊𑆋𑆌𑆍𑆎𑆏𑆐𑆑𑆒𑆓𑆔𑆕𑆖𑆗𑆘𑆙𑆚𑆛𑆜𑆝𑆞𑆟𑆠𑆡𑆢𑆣𑆤𑆥𑆦𑆧𑆨𑆩𑆪𑆫𑆬𑆭𑆮𑆯𑆰𑆱𑆲𑆳𑆴𑆵𑆶𑆷𑆸𑆹𑆺𑆻𑆼𑆽𑆾𑆿𑇀𑇁𑇂𑇃𑇄𑇅𑇆𑇇𑇈𑇉𑇊𑇋𑇌𑇍𑇎𑇏𑇐𑇑𑇒𑇓𑇔𑇕𑇖𑇗𑇘𑇙𑇚𑇛𑇜𑇝𑇞𑇟𑇠𑇡𑇢𑇣𑇤𑇥𑇦𑇧𑇨𑇩𑇪𑇫𑇬𑇭𑇮𑇯𑇰𑇱𑇲𑇳𑇴𑇵𑇶𑇷𑇸𑇹𑇺𑇻𑇼𑇽𑇾𑇿𑈀𑈁𑈂𑈃𑈄𑈅𑈆𑈇𑈈𑈉𑈊𑈋𑈌𑈍𑈎𑈏𑈐𑈑𑈒𑈓𑈔𑈕𑈖𑈗𑈘𑈙𑈚𑈛𑈜𑈝𑈞𑈟𑈠𑈡𑈢𑈣𑈤𑈥𑈦𑈧𑈨𑈩𑈪𑈫𑈬𑈭𑈮𑈯𑈰𑈱𑈲𑈳𑈴𑈶𑈵𑈷𑈸𑈹𑈺𑈻𑈼𑈽𑈾𑈿𑉀𑉁𑉂𑉃𑉄𑉅𑉆𑉇𑉈𑉉𑉊𑉋𑉌𑉍𑉎𑉏𑉐𑉑𑉒𑉓𑉔𑉕𑉖𑉗𑉘𑉙𑉚𑉛𑉜𑉝𑉞𑉟𑉠𑉡𑉢𑉣𑉤𑉥𑉦𑉧𑉨𑉩𑉪𑉫𑉬𑉭𑉮𑉯𑉰𑉱𑉲𑉳𑉴𑉵𑉶𑉷𑉸𑉹𑉺𑉻𑉼𑉽𑉾𑉿𑊀𑊁𑊂𑊃𑊄𑊅𑊆𑊇𑊈𑊉𑊊𑊋𑊌𑊍𑊎𑊏𑊐𑊑𑊒𑊓𑊔𑊕𑊖𑊗𑊘𑊙𑊚𑊛𑊜𑊝𑊞𑊟𑊠𑊡𑊢𑊣𑊤𑊥𑊦𑊧𑊨𑊩𑊪𑊫𑊬𑊭𑊮𑊯𑊰𑊱𑊲𑊳𑊴𑊵𑊶𑊷𑊸𑊹𑊺𑊻𑊼𑊽𑊾𑊿𑋀𑋁𑋂𑋃𑋄𑋅𑋆𑋇𑋈𑋉𑋊𑋋𑋌𑋍𑋎𑋏𑋐𑋑𑋒𑋓𑋔𑋕𑋖𑋗𑋘𑋙𑋚𑋛𑋜𑋝𑋞𑋟𑋠𑋡𑋢𑋣𑋤𑋥𑋦𑋧𑋨𑋩𑋪𑋫𑋬𑋭𑋮𑋯𑋰𑋱𑋲𑋳𑋴𑋵𑋶𑋷𑋸𑋹𑋺𑋻𑋼𑋽𑋾𑋿𑌀𑌁𑌂𑌃𑌄𑌅𑌆𑌇𑌈𑌉𑌊𑌋𑌌𑌍𑌎𑌏𑌐𑌑𑌒𑌓𑌔𑌕𑌖𑌗𑌘𑌙𑌚𑌛𑌜𑌝𑌞𑌟𑌠𑌡𑌢𑌣𑌤𑌥𑌦𑌧𑌨𑌩𑌪𑌫𑌬𑌭𑌮𑌯𑌰𑌱𑌲𑌳𑌴𑌵𑌶𑌷𑌸𑌹𑌺𑌻𑌼𑌽𑌾𑌿𑍀𑍁𑍂𑍃𑍄𑍅𑍆𑍇𑍈𑍉𑍊𑍋𑍌𑍍𑍎𑍏𑍐𑍑𑍒𑍓𑍔𑍕𑍖𑍗𑍘𑍙𑍚𑍛𑍜𑍝𑍞𑍟𑍠𑍡𑍢𑍣𑍤𑍥𑍦𑍧𑍨𑍩𑍪𑍫𑍬𑍭𑍮𑍯𑍰𑍱𑍲𑍳𑍴𑍵𑍶𑍷𑍸𑍹𑍺𑍻𑍼𑍽𑍾𑍿𑎀𑎁𑎂𑎃𑎄𑎅𑎆𑎇𑎈𑎉𑎊𑎋𑎌𑎍𑎎𑎏𑎐𑎑𑎒𑎓𑎔𑎕𑎖𑎗𑎘𑎙𑎚𑎛𑎜𑎝𑎞𑎟𑎠𑎡𑎢𑎣𑎤𑎥𑎦𑎧𑎨𑎩𑎪𑎫𑎬𑎭𑎮𑎯𑎰𑎱𑎲𑎳𑎴𑎵𑎶𑎷𑎸𑎹𑎺𑎻𑎼𑎽𑎾𑎿𑏀𑏁𑏂𑏃𑏄𑏅𑏆𑏇𑏈𑏉𑏊𑏋𑏌𑏍𑏎𑏏𑏐𑏑𑏒𑏓𑏔𑏕𑏖𑏗𑏘𑏙𑏚𑏛𑏜𑏝𑏞𑏟𑏠𑏡𑏢𑏣𑏤𑏥𑏦𑏧𑏨𑏩𑏪𑏫𑏬𑏭𑏮𑏯𑏰𑏱𑏲𑏳𑏴𑏵𑏶𑏷𑏸𑏹𑏺𑏻𑏼𑏽𑏾𑏿𑐀𑐁𑐂𑐃𑐄𑐅𑐆𑐇𑐈𑐉𑐊𑐋𑐌𑐍𑐎𑐏𑐐𑐑𑐒𑐓𑐔𑐕𑐖𑐗𑐘𑐙𑐚𑐛𑐜𑐝𑐞𑐟𑐠𑐡𑐢𑐣𑐤𑐥𑐦𑐧𑐨𑐩𑐪𑐫𑐬𑐭𑐮𑐯𑐰𑐱𑐲𑐳𑐴𑐵𑐶𑐷𑐸𑐹𑐺𑐻𑐼𑐽𑐾𑐿𑑀𑑁𑑂𑑃𑑄𑑅𑑆𑑇𑑈𑑉𑑊𑑋𑑌𑑍𑑎𑑏𑑐𑑑𑑒𑑓𑑔𑑕𑑖𑑗𑑘𑑙𑑚𑑛𑑜𑑝𑑞𑑟𑑠𑑡𑑢𑑣𑑤𑑥𑑦𑑧𑑨𑑩𑑪𑑫𑑬𑑭𑑮𑑯𑑰𑑱𑑲𑑳𑑴𑑵𑑶𑑷𑑸𑑹𑑺𑑻𑑼𑑽𑑾𑑿𑒀𑒁𑒂𑒃𑒄𑒅𑒆𑒇𑒈𑒉𑒊𑒋𑒌𑒍𑒎𑒏𑒐𑒑𑒒𑒓𑒔𑒕𑒖𑒗𑒘𑒙𑒚𑒛𑒜𑒝𑒞𑒟𑒠𑒡𑒢𑒣𑒤𑒥𑒦𑒧𑒨𑒩𑒪𑒫𑒬𑒭𑒮𑒯𑒰𑒱𑒲𑒳𑒴𑒵𑒶𑒷𑒸𑒻𑒻𑒼𑒽𑒾𑒿𑓀𑓁𑓃𑓂𑓄𑓅𑓆𑓇𑓈𑓉𑓊𑓋𑓌𑓍𑓎𑓏𑓐𑓑𑓒𑓓𑓔𑓕𑓖𑓗𑓘𑓙𑓚𑓛𑓜𑓝𑓞𑓟𑓠𑓡𑓢𑓣𑓤𑓥𑓦𑓧𑓨𑓩𑓪𑓫𑓬𑓭𑓮𑓯𑓰𑓱𑓲𑓳𑓴𑓵𑓶𑓷𑓸𑓹𑓺𑓻𑓼𑓽𑓾𑓿𑔀𑔁𑔂𑔃𑔄𑔅𑔆𑔇𑔈𑔉𑔊𑔋𑔌𑔍𑔎𑔏𑔐𑔑𑔒𑔓𑔔𑔕𑔖𑔗𑔘𑔙𑔚𑔛𑔜𑔝𑔞𑔟𑔠𑔡𑔢𑔣𑔤𑔥𑔦𑔧𑔨𑔩𑔪𑔫𑔬𑔭𑔮𑔯𑔰𑔱𑔲𑔳𑔴𑔵𑔶𑔷𑔸𑔹𑔺𑔻𑔼𑔽𑔾𑔿𑕀𑕁𑕂𑕃𑕄𑕅𑕆𑕇𑕈𑕉𑕊𑕋𑕌𑕍𑕎𑕏𑕐𑕑𑕒𑕓𑕔𑕕𑕖𑕗𑕘𑕙𑕚𑕛𑕜𑕝𑕞𑕟𑕠𑕡𑕢𑕣𑕤𑕥𑕦𑕧𑕨𑕩𑕪𑕫𑕬𑕭𑕮𑕯𑕰𑕱𑕲𑕳𑕴𑕵𑕶𑕷𑕸𑕹𑕺𑕻𑕼𑕽𑕾𑕿𑖀𑖁𑖂𑖃𑖄𑖅𑖆𑖇𑖈𑖉𑖊𑖋𑖌𑖍𑖎𑖏𑖐𑖑𑖒𑖓𑖔𑖕𑖖𑖗𑖘𑖙𑖚𑖛𑖜𑖝𑖞𑖟𑖠𑖡𑖢𑖣𑖤𑖥𑖦𑖧𑖨𑖩𑖪𑖫𑖬𑖭𑖮𑖯𑖰𑖱𑖲𑖳𑖴𑖵𑖶𑖷𑖸𑖹𑖺𑖻𑖼𑖽𑖾𑗀𑖿𑗁𑗂𑗃𑗄𑗅𑗆𑗇𑗈𑗉𑗊𑗋𑗌𑗍𑗎𑗏𑗐𑗑𑗒𑗓𑗔𑗕𑗖𑗗𑗘𑗙𑗚𑗛𑗜𑗝𑗞𑗟𑗠𑗡𑗢𑗣𑗤𑗥𑗦𑗧𑗨𑗩𑗪𑗫𑗬𑗭𑗮𑗯𑗰𑗱𑗲𑗳𑗴𑗵𑗶𑗷𑗸𑗹𑗺𑗻𑗼𑗽𑗾𑗿𑘀𑘁𑘂𑘃𑘄𑘅𑘆𑘇𑘈𑘉𑘊𑘋𑘌𑘍𑘎𑘏𑘐𑘑𑘒𑘓𑘔𑘕𑘖𑘗𑘘𑘙𑘚𑘛𑘜𑘝𑘞𑘟𑘠𑘡𑘢𑘣𑘤𑘥𑘦𑘧𑘨𑘩𑘪𑘫𑘬𑘭𑘮𑘯𑘰𑘱𑘲𑘳𑘴𑘵𑘶𑘷𑘸𑘹𑘺𑘻𑘼𑘽𑘾𑘿𑙀𑙁𑙂𑙃𑙄𑙅𑙆𑙇𑙈𑙉𑙊𑙋𑙌𑙍𑙎𑙏𑙐𑙑𑙒𑙓𑙔𑙕𑙖𑙗𑙘𑙙𑙚𑙛𑙜𑙝𑙞𑙟𑙠𑙡𑙢𑙣𑙤𑙥𑙦𑙧𑙨𑙩𑙪𑙫𑙬𑙭𑙮𑙯𑙰𑙱𑙲𑙳𑙴𑙵𑙶𑙷𑙸𑙹𑙺𑙻𑙼𑙽𑙾𑙿𑚀𑚁𑚂𑚃𑚄𑚅𑚆𑚇𑚈𑚉𑚊𑚋𑚌𑚍𑚎𑚏𑚐𑚑𑚒𑚓𑚔𑚕𑚖𑚗𑚘𑚙𑚚𑚛𑚜𑚝𑚞𑚟𑚠𑚡𑚢𑚣𑚤𑚥𑚦𑚧𑚨𑚩𑚪𑚫𑚬𑚭𑚮𑚯𑚰𑚱𑚲𑚳𑚴𑚵𑚷𑚶𑚸𑚹𑚺𑚻𑚼𑚽𑚾𑚿𑛀𑛁𑛂𑛃𑛄𑛅𑛆𑛇𑛈𑛉𑛊𑛋𑛌𑛍𑛎𑛏𑛐𑛑𑛒𑛓𑛔𑛕𑛖𑛗𑛘𑛙𑛚𑛛𑛜𑛝𑛞𑛟𑛠𑛡𑛢𑛣𑛤𑛥𑛦𑛧𑛨𑛩𑛪𑛫𑛬𑛭𑛮𑛯𑛰𑛱𑛲𑛳𑛴𑛵𑛶𑛷𑛸𑛹𑛺𑛻𑛼𑛽𑛾𑛿𑜀𑜁𑜂𑜃𑜄𑜅𑜆𑜇𑜈𑜉𑜊𑜋𑜌𑜍𑜎𑜏𑜐𑜑𑜒𑜓𑜔𑜕𑜖𑜗𑜘𑜙𑜚𑜛𑜜𑜝𑜞𑜟𑜠𑜡𑜢𑜣𑜤𑜥𑜦𑜧𑜨𑜩𑜪𑜫𑜬𑜭𑜮𑜯𑜰𑜱𑜲𑜳𑜴𑜵𑜶𑜷𑜸𑜹𑜺𑜻𑜼𑜽𑜾𑜿𑝀𑝁𑝂𑝃𑝄𑝅𑝆𑝇𑝈𑝉𑝊𑝋𑝌𑝍𑝎𑝏𑝐𑝑𑝒𑝓𑝔𑝕𑝖𑝗𑝘𑝙𑝚𑝛𑝜𑝝𑝞𑝟𑝠𑝡𑝢𑝣𑝤𑝥𑝦𑝧𑝨𑝩𑝪𑝫𑝬𑝭𑝮𑝯𑝰𑝱𑝲𑝳𑝴𑝵𑝶𑝷𑝸𑝹𑝺𑝻𑝼𑝽𑝾𑝿𑞀𑞁𑞂𑞃𑞄𑞅𑞆𑞇𑞈𑞉𑞊𑞋𑞌𑞍𑞎𑞏𑞐𑞑𑞒𑞓𑞔𑞕𑞖𑞗𑞘𑞙𑞚𑞛𑞜𑞝𑞞𑞟𑞠𑞡𑞢𑞣𑞤𑞥𑞦𑞧𑞨𑞩𑞪𑞫𑞬𑞭𑞮𑞯𑞰𑞱𑞲𑞳𑞴𑞵𑞶𑞷𑞸𑞹𑞺𑞻𑞼𑞽𑞾𑞿𑟀𑟁𑟂𑟃𑟄𑟅𑟆𑟇𑟈𑟉𑟊𑟋𑟌𑟍𑟎𑟏𑟐𑟑𑟒𑟓𑟔𑟕𑟖𑟗𑟘𑟙𑟚𑟛𑟜𑟝𑟞𑟟𑟠𑟡𑟢𑟣𑟤𑟥𑟦𑟧𑟨𑟩𑟪𑟫𑟬𑟭𑟮𑟯𑟰𑟱𑟲𑟳𑟴𑟵𑟶𑟷𑟸𑟹𑟺𑟻𑟼𑟽𑟾𑟿𑠀𑠁𑠂𑠃𑠄𑠅𑠆𑠇𑠈𑠉𑠊𑠋𑠌𑠍𑠎𑠏𑠐𑠑𑠒𑠓𑠔𑠕𑠖𑠗𑠘𑠙𑠚𑠛𑠜𑠝𑠞𑠟𑠠𑠡𑠢𑠣𑠤𑠥𑠦𑠧𑠨𑠩𑠪𑠫𑠬𑠭𑠮𑠯𑠰𑠱𑠲𑠳𑠴𑠵𑠶𑠷𑠸𑠺𑠹𑠻𑠼𑠽𑠾𑠿𑡀𑡁𑡂𑡃𑡄𑡅𑡆𑡇𑡈𑡉𑡊𑡋𑡌𑡍𑡎𑡏𑡐𑡑𑡒𑡓𑡔𑡕𑡖𑡗𑡘𑡙𑡚𑡛𑡜𑡝𑡞𑡟𑡠𑡡𑡢𑡣𑡤𑡥𑡦𑡧𑡨𑡩𑡪𑡫𑡬𑡭𑡮𑡯𑡰𑡱𑡲𑡳𑡴𑡵𑡶𑡷𑡸𑡹𑡺𑡻𑡼𑡽𑡾𑡿𑢀𑢁𑢂𑢃𑢄𑢅𑢆𑢇𑢈𑢉𑢊𑢋𑢌𑢍𑢎𑢏𑢐𑢑𑢒𑢓𑢔𑢕𑢖𑢗𑢘𑢙𑢚𑢛𑢜𑢝𑢞𑢟𑢠𑢡𑢢𑢣𑢤𑢥𑢦𑢧𑢨𑢩𑢪𑢫𑢬𑢭𑢮𑢯𑢰𑢱𑢲𑢳𑢴𑢵𑢶𑢷𑢸𑢹𑢺𑢻𑢼𑢽𑢾𑢿𑣀𑣁𑣂𑣃𑣄𑣅𑣆𑣇𑣈𑣉𑣊𑣋𑣌𑣍𑣎𑣏𑣐𑣑𑣒𑣓𑣔𑣕𑣖𑣗𑣘𑣙𑣚𑣛𑣜𑣝𑣞𑣟𑣠𑣡𑣢𑣣𑣤𑣥𑣦𑣧𑣨𑣩𑣪𑣫𑣬𑣭𑣮𑣯𑣰𑣱𑣲𑣳𑣴𑣵𑣶𑣷𑣸𑣹𑣺𑣻𑣼𑣽𑣾𑣿𑤀𑤁𑤂𑤃𑤄𑤅𑤆𑤇𑤈𑤉𑤊𑤋𑤌𑤍𑤎𑤏𑤐𑤑𑤒𑤓𑤔𑤕𑤖𑤗𑤘𑤙𑤚𑤛𑤜𑤝𑤞𑤟𑤠𑤡𑤢𑤣𑤤𑤥𑤦𑤧𑤨𑤩𑤪𑤫𑤬𑤭𑤮𑤯𑤰𑤱𑤲𑤳𑤴𑤵𑤶𑤷𑤸𑤹𑤺𑤻𑤼𑤽𑤾𑤿𑥀𑥁𑥂𑥃𑥄𑥅𑥆𑥇𑥈𑥉𑥊𑥋𑥌𑥍𑥎𑥏𑥐𑥑𑥒𑥓𑥔𑥕𑥖𑥗𑥘𑥙𑥚𑥛𑥜𑥝𑥞𑥟𑥠𑥡𑥢𑥣𑥤𑥥𑥦𑥧𑥨𑥩𑥪𑥫𑥬𑥭𑥮𑥯𑥰𑥱𑥲𑥳𑥴𑥵𑥶𑥷𑥸𑥹𑥺𑥻𑥼𑥽𑥾𑥿𑦀𑦁𑦂𑦃𑦄𑦅𑦆𑦇𑦈𑦉𑦊𑦋𑦌𑦍𑦎𑦏𑦐𑦑𑦒𑦓𑦔𑦕𑦖𑦗𑦘𑦙𑦚𑦛𑦜𑦝𑦞𑦟𑦠𑦡𑦢𑦣𑦤𑦥𑦦𑦧𑦨𑦩𑦪𑦫𑦬𑦭𑦮𑦯𑦰𑦱𑦲𑦳𑦴𑦵𑦶𑦷𑦸𑦹𑦺𑦻𑦼𑦽𑦾𑦿𑧀𑧁𑧂𑧃𑧄𑧅𑧆𑧇𑧈𑧉𑧊𑧋𑧌𑧍𑧎𑧏𑧐𑧑𑧒𑧓𑧔𑧕𑧖𑧗𑧘𑧙𑧚𑧛𑧜𑧝𑧞𑧟𑧠𑧡𑧢𑧣𑧤𑧥𑧦𑧧𑧨𑧩𑧪𑧫𑧬𑧭𑧮𑧯𑧰𑧱𑧲𑧳𑧴𑧵𑧶𑧷𑧸𑧹𑧺𑧻𑧼𑧽𑧾𑧿𑨀𑨁𑨂𑨃𑨄𑨅𑨆𑨇𑨈𑨉𑨊𑨋𑨌𑨍𑨎𑨏𑨐𑨑𑨒𑨓𑨔𑨕𑨖𑨗𑨘𑨙𑨚𑨛𑨜𑨝𑨞𑨟𑨠𑨡𑨢𑨣𑨤𑨥𑨦𑨧𑨨𑨩𑨪𑨫𑨬𑨭𑨮𑨯𑨰𑨱𑨲𑨳𑨴𑨵𑨶𑨷𑨸𑨹𑨺𑨻𑨼𑨽𑨾𑨿𑩀𑩁𑩂𑩃𑩄𑩅𑩆𑩇𑩈𑩉𑩊𑩋𑩌𑩍𑩎𑩏𑩐𑩑𑩒𑩓𑩔𑩕𑩖𑩗𑩘𑩙𑩚𑩛𑩜𑩝𑩞𑩟𑩠𑩡𑩢𑩣𑩤𑩥𑩦𑩧𑩨𑩩𑩪𑩫𑩬𑩭𑩮𑩯𑩰𑩱𑩲𑩳𑩴𑩵𑩶𑩷𑩸𑩹𑩺𑩻𑩼𑩽𑩾𑩿𑪀𑪁𑪂𑪃𑪄𑪅𑪆𑪇𑪈𑪉𑪊𑪋𑪌𑪍𑪎𑪏𑪐𑪑𑪒𑪓𑪔𑪕𑪖𑪗𑪘𑪙𑪚𑪛𑪜𑪝𑪞𑪟𑪠𑪡𑪢𑪣𑪤𑪥𑪦𑪧𑪨𑪩𑪪𑪫𑪬𑪭𑪮𑪯𑪰𑪱𑪲𑪳𑪴𑪵𑪶𑪷𑪸𑪹𑪺𑪻𑪼𑪽𑪾𑪿𑫀𑫁𑫂𑫃𑫄𑫅𑫆𑫇𑫈𑫉𑫊𑫋𑫌𑫍𑫎𑫏𑫐𑫑𑫒𑫓𑫔𑫕𑫖𑫗𑫘𑫙𑫚𑫛𑫜𑫝𑫞𑫟𑫠𑫡𑫢𑫣𑫤𑫥𑫦𑫧𑫨𑫩𑫪𑫫𑫬𑫭𑫮𑫯𑫰𑫱𑫲𑫳𑫴𑫵𑫶𑫷𑫸𑫹𑫺𑫻𑫼𑫽𑫾𑫿𑬀𑬁𑬂𑬃𑬄𑬅𑬆𑬇𑬈𑬉𑬊𑬋𑬌𑬍𑬎𑬏𑬐𑬑𑬒𑬓𑬔

between consonants of similar manner, and is more related to the general pattern of manner assimilation observed in MIA where the *l* is characterised as stronger than *y* (§ 6.1.3). Other possibilities include the transcription of borrowed MIA forms, or orthographic laxity in relation to the awkward flourish of the *y* diacritic (§ 8.3.1.4.3). While the example of Skt. *kalyāṇa-* > G *kalāṇa-* from the Aśokan inscriptions suggests the reduced form, whether phonetic or orthographic, was in circulation at an early stage, the *ly* cluster *akṣara* is also recorded in the third century CE Schøyen manuscripts (Glass 2000: 121), although it is considered to be a restored “Sanskritised” form in this context (§ 8.7).

8.3.4.1.5 Summary of *l* + C cluster developments

A summary of the more widely attested Gāndhārī reflexes of OIA *l* + C clusters is presented in Table 8.3.4.1.5.

Table 8.3.4.1.5: Summary of *l* + C cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/-lp-/	/-pp-/	[-pp-]	/-pp-/ ≡ {/-p-/ ^{adv} , (/rp-/ ^{alt/h})} The geminate is typically represented by a single stop, while /rp-/ ^{alt/h} may be an alternate marking of either the geminate or the retained cluster.
/-ly-/	/-ll-/	[-ll-]	/-ll-/ ≡ {/-l-/ ^{adv} , (/ly-/ ^h)} The geminate <i>l</i> is typically represented by a single <i>akṣara</i> , while /ly-/ ^h is an historic form recorded in proper nouns and some later manuscripts.

8.3.4.2 *ν* + C clusters

8.3.4.2.1 *ν* + stop clusters

There are no *ν* + stop clusters recorded in OIA.

8.3.4.2.2 *ν* + nasal clusters

There are only two *ν* + nasal clusters recorded in OIA, *νn* and *νṇ*, neither of which is recorded in the Gāndhārī source materials.

8.3.4.2.3 ν + fricative clusters

There is only one ν + fricative clusters recorded in OIA, νh , which is not recorded in the Gāndhārī source materials.

8.3.4.2.4 ν + semi-vowel clusters

Examples: Gāndhārī reflexes of OIA ν + semi-vowel clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.4.2.4a.

Table 8.3.4.2.4a: Gāndhārī reflexes of ν + semi-vowel clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
νy -	ν -	Aś-M	$\nu yañjana$ - > $\nu añaana$ -; typically reduced	3 BCE
		BLS4-1	$\nu yāpāda$ - > $\nu a[va]da$ -	1 CE
		BLS9-1	$\nu yapanudya$ > $\nu aṇuja$	
	νi -	BLS1-1	$\nu yākārṣam$ > $\nu iaghaṣe$	1 CE
		BLS1-2	BHS $\nu yavasarga$ - > $\nu ivasagha$ -	
		BLS9-1	$\nu yañjanāni$ > $\nu iṃjaṇaṇi$	
		RS-14	$\nu yāda$ - > $\nu iāda$ -	2 CE
		BCS5-1	$\nu yatimiśra$ - > $\nu idimiśa$ -	2 CE
		BCS5-2	$\nu yārtha(ka)$ - > $\nu iārtha(e)$ -	
νy -	$\nu(v)$ -	Aś-M	$kartavya$ - > $kartava$ -; typically reduced	3 BCE
		InsK-M	$vāstavya$ - > $vastava$ -	1 BCE
			$vāstavya$ - > $\nu vastava$ -	2 CE
		BLS1-1	$gurukartavyaḥ$ > $ghuruhatavu$	1 CE
		BLS1-2	$anuvyañjana$ - > $aṇovejaṇa$ -	
		BLS2-2	$mantritavyam$ > $matridavo$	
		BLS4-1	$jñātavyam$ > $ñadava$, $divyayā$ > $[di]vae$	
		RS-19	$spraṣṭavya$ - > $phaṭhava$ -	2 CE
		BCS5-1	$\nu tavya$ - > $\nu dava$ -	
		BLS1-2	BHS $spraṣṭavya$ - > $proṭhabu$ -	1 CE
νr -	νr -	Aś-M	$\nu raja$ - > $\nu raca$ -; typically stable	3 BCE
	ν -	BLS2-2	$\nu rajati$ > $\nu acadi$; uncertain	1 CE
νr -	νr -	Aś-M	$pravrajita$ - > $pravrajita$ -; typically stable	3 BCE

	-v(v)-	BLS4-1	<i>parivrajet > parivraye</i>	1 CE
		BLS1-1	<i>pravrajyām > pravaja</i> ; possibly borrowed	1 CE
		BLS2-1	<i>pravrājitaḥ > pravayido</i> ; possibly borrowed	
		BLS4-1	<i>parivrajet > parivaye</i>	
	-b(b)-	RS-19	<i>pravrajyām > pravace</i>	2 CE
		BLS1-1	BHS <i>*pravrājesi > prabaeṣi</i>	1 CE
		BLS1-2	<i>āvrajeyam > [abaji]</i> ; uncertain	

Other OIA combinations not recorded in the Gāndhārī source documents are *vv* and *vl*.

Overview: In word initial positions the reflex of *vy-* is recorded as *vi-* in a number of cases, while it is also reduced to *v-* in others, while in medial positions *-vy-* is assimilated to *-v(v)-* in all but one case where it is recorded as *-b(b)-*. While *vy* has been recorded using the typical wavy line diacritic (§ 5.3.2.1) in the Schøyen manuscripts (Glass 2000: 121), there are no instances of this usage in the selected source materials²⁷³. In a more limited sample word initial *vr-* appears to remain stable in the Aśokan inscriptions, while being reduced to *v-* in one uncertain reading from the BL manuscripts. Medial *-vr-* is mostly assimilated to *-v(v)-* or *-b(b)-*, while also being retained in the Aśokan inscriptions and one example from the BL manuscripts. Post-consonantal *r* is indicated in the usual way in the Kharoṣṭhī script by an underscore diacritic stroke (§ 5.3.2.1).

Discussion: In word initial position the predominant reflex of the *vy-* cluster is *vi-*, although the *y* is also deleted in some cases. Since the phonetic features and acoustic properties of *y* and *i* are similar (Kent and Read 2002: 177-179) the writing of *vi-* may in some cases represent a graphic shorthand in which the simpler form of *vi* (𑀧) is used, compared with the more complex *vy* (𑀧𑀭) graphic in the Kharoṣṭhī script (§ 5.3.2.1). Assuming there was no intention to adjust the syllable structure in the sense of Skt. *vyā.da-* > G *via.da-* (RS-14), compared with say *vi.a.da-*, this suggests a retention of the initial *vy-* cluster in many cases, at least in a phonetic sense. The alternate reflexes of *v-* suggest that *y* may have been reduced through lenition and deletion, which is akin to the MIA development as *b-* (§ 9.1) and raises the possibility that these were borrowed forms.

²⁷³ As noted in § 5.3.2.1, simpler forms of diacritic marked characters for *vya* and *vyo* have been observed in preliminary studies of the BL fragment 4 by Tim Lenz. These are used in Gāndhārī words which have been interpreted as *vyadharado* and *abhavyo* respectively (T.Lenz, personal communication, September 11 and 18, 2023).

In medial positions -vy- appears to have been assimilated to geminate -vv- in all cases except one, which is consistent with the typical Gāndhārī pattern of assimilating *y* in medial Ty, Ny and Sy clusters. Since single intervocalic -v- is typically preserved in Gāndhārī (§ 8.2.4.2), the assumption of geminate -v(v)- reflexes relies on parallel MIA developments of -vv- and -bb- (§ 9.2.2). The labial stop -b(b)- reflex supports this assumption, since single OIA intervocalic -b- is often weakened to -v-, and the writing of unweakened -b- normally indicates gemination.

The word initial *vr-* cluster appears to have been retained in the third century BCE Aśokan inscriptions, while it was possibly reduced by the time of the first century BL manuscripts, although this is based on an uncertain example. The medial -*vr-* clusters are typically stable in the Aśokan inscriptions²⁷⁴, but appear to be assimilated to the -v(v)- geminate in the BL and RS manuscript examples. This suggest a variable record, although the latter use occurs in Buddhist technical terms derived from the Sanskrit verbal root √*vraj*, and may be borrowings from the MIA substrate dialect, rather than a true representation of the Gāndhārī phonetic development²⁷⁵. The retention of post-consonantal *r* in -*vr-* is consistent with the common Gāndhārī pattern seen in the *Tr* and *Nr* clusters discussed in previous sections, although this occurs less reliably in the case of *Sr* clusters.

8.3.4.2.5 Summary of *v* + C cluster developments

A summary of the Gāndhārī reflexes of OIA *v* + C clusters is presented in Table 8.3.4.2.5. A larger data base is required to confirm the developments interpreted in the above discussion, and conservative assumptions have been made in the tabulated reflexes.

Table 8.3.4.2.5: Summary of *v* + C cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
/vy-/	/vy-/	[vj / vi]	/vy-/ ≡ {/vi-/ ^h , (/v-/ ^{adv})} The retained cluster is represented by the allophone /vi-/ ^h while the advanced form /v-/ ^{adv} is also used.

²⁷⁴ Hultzsch 1925: lxxxix; Mehendale 1948: xx.

²⁷⁵ Lenz 2003: 132; Salomon 2008: 121.

/-vy-/	/-vv-/	[-vv-]	/-vv-/ \equiv {/-v-/ ^{adv} , (/b-/ ^{alt})} The geminate is typically represented by a single stop, while /b-/ ^{alt} is an alternate labial development.
/vr-/	/vr-/	[vɪ-]	/vr-/ \equiv {/vr-/ ^h , /v-/ ^{adv} } The cluster is represented by both the historic /vr-/ ^h and advanced forms /v-/ ^{adv} in a limited sample.
/-vr-/	/-vr-/	[-vɪ-]	/-vr-/ \equiv {/-vr-/ ^h , /v-/ ^{alt} , (/b-/ ^{alt})} The cluster is represented by the historic form /-vr-/ ^h , while /v-/ ^{alt} and /b-/ ^{alt} may be borrowings.

8.3.4.3 y + R clusters

There are no yT, yN or yS clusters recorded in OIA, while the only yR clusters are yv and yy. Of these, only the geminate yy is recorded in the Gāndhārī source documents in BHS *sayyathāpi* > G *ṣayaṣavi*, seen in RS-5 and RS-20 (2 CE), where it appears to remain stable. This is represented in Table 8.3.4.3.

Table 8.3.4.3: Summary of y + R cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes Typical reflexes are treated as advanced forms (adv).
/-yy-/	/-yy-/	[jj]	/-yy-/ \equiv /-y-/ ^{adv} The geminate is represented by a single <i>akṣara</i> .

8.3.4.4 r + C clusters

OIA clusters involving pre-consonantal *r* are very common, occurring in -rT-, -rN-, -rS- and -rR- clusters, where *r* is combined with consonants having all manners of articulation. They only occur word medially, and Gāndhārī reflexes are recorded for all of the manner combinations. In many cases these clusters are preserved, and the preconsonantal *r* is typically indicated in the Kharoṣṭhī script using different forms of diacritic attached to the main *akṣara* (§ 5.3.2.2). These vary from a central horizontal strike through the stem seen in the Aśokan inscriptions, as in *rva* (𑀢𑀸) for example, through variants on a clockwise curved line attached to the base of the *akṣara* seen in the manuscript sources, as in *rga* (𑀢𑀸𑀓) for example. These are all distinct from the rightward horizontal underscore diacritic used to mark post-consonantal *r* (§ 5.3.2.1).

As discussed in § 8.3.1.4.4, it is possible that the OIA retroflex tap *r* was realised as an alveolar continuant such as [ɹ]. It was suggested that a weakened articulation such as this could also occur in the *Tr* consonant clusters, where a slightly more open articulation of the rhotic may have occurred in the transition from the stop to the following vowel. A similar more open articulation of the rhotic is also possible during closure to the obstruent in the *rT* and *rN* clusters.

8.3.4.4.1 *r* + stop clusters

Examples: Gāndhārī reflexes of OIA *r* + stop clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.4.4.1a.

Table 8.3.4.4.1a: Gāndhārī reflexes of *r* + stop clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes (1) -Cr- considered orthographic (Hultzsch 1925: lxxvii)		Era
-rk-	-rk-	RS-12	<i>avitarkam</i> > <i>avitarka</i>	2 CE
	-(k)k-	BLS4-1	-vitarka- > -vitra[ka]-	1 CE
-rg-	-rg-	InsK-M	<i>Mārgaśiras-</i> > <i>Ma[rgaśi]ras(r)a-</i>	3 CE
		BLS2-1	<i>mārgaṇa</i> > <i>marga[n]o</i>	1 CE
		BCS5-1	<i>durgati</i> > <i>durgadi-</i>	2 CE
	-rVg-	BCS5-1	<i>durga-</i> > <i>droga-</i> , <i>dru[ga]-</i>	2 CE
	-gr- (1)	Aś-H	<i>varga-</i> > <i>vagra-</i> , <i>svarga-</i> > <i>spagra-</i>	3 BCE
	-(g)g-	BLS2-2	<i>saṅghamarga-</i> > <i>saghamag[o]-</i>	1 CE
		BLS4-1	<i>mārga-</i> > <i>maga-</i>	
		BLS9-1	-mārgaḥ > -mago	
		RS-5	<i>mārgasya</i> > <i>magaṣa</i>	2 CE
	-g-	RS-12	BHS <i>mārgātaḥ</i> > <i>maga(*de)</i>	
		RS-12	<i>mārgaṇa</i> > <i>magaṇa</i>	2 CE
		BLS1-2	<i>mārga-</i> > <i>magha-</i> , BHS <i>vyavasarga-</i> > <i>vivasagha-</i> (gg written as gh)	1 CE
	-gh-			
-rgh-	-rgh-	InsK-M	<i>dīrghāyus-</i> > <i>dirghayu[ta*]</i>	3 CE
	-(g)gh-	InsK-M	<i>argha-</i> > <i>agha-</i>	1 BCE
	[-g-]	RS-12	<i>dīrgha-</i> > <i>driga-</i>	2 CE
	[-kh-]	RS-12	<i>dīrgha-</i> > <i>drikha-</i>	2 CE

-rc-	-(c)c-	BLS1-2	<i>arciṣmanti</i> > <i>aceata</i>	1 CE
-rj-	-rj-	BLS4-1	<i>vivarj</i> ayitvā > <i>vivarjita</i>	1 CE
	-(j)j-	BLS9-1	<i>vivarj</i> ayitvā > <i>viva</i> jaita	1 CE
		RS-12	<i>saṃmrjya-</i> > <i>samaja-</i>	2 CE
-rt-	-rt-	BLS9-1	secondary: <i>vṛtti-</i> > * <i>vruti-</i> > <i>vurti-</i> (?)	1 CE
		BLS4-1	In <i>vartamaṇi-</i> (?)	
	-tr- (1)	Aś-H	<i>kirti-</i> > <i>kitri-</i>	3 BCE
	-(t)t-	Aś-H	<i>anuvarta-</i> > <i>anuvata-</i> , <i>nirvarta-</i> > <i>nivata-</i>	3 BCE
		BLS1-1	* <i>prāvartīṣam</i> > <i>pravatiṣu</i>	1 CE
		BLS1-2	<i>kartum</i> > <i>katu/kato</i>	
		BLS4-1	- <i>vartani-</i> > - <i>vataṇi-</i>	
		RS-12	<i>muhūrtam</i> > <i>mahuta</i>	2 CE
		RS-19	<i>āvartate</i> > <i>avataḍi</i>	
		RS-20	<i>saṃvartanīyān</i> > - <i>ṣabataṇia</i>	
	-(t)ṭ-	BLS1-1	* <i>varta-</i> > <i>baṭa-</i>	1 CE
		RS-19	<i>āvarta-</i> > <i>avaṭa-</i>	2 CE
		BCS5-1	<i>kartavyam</i> > <i>kaṭave</i>	2 CE
	-(ṭ)ṭh-	RS-12	- <i>vartataḥ</i> > <i>vaṭhaḍa</i>	2 CE
-rth-	-rth-	Aś-H	<i>artha-</i> > <i>artha-</i>	3 BCE
		BLS2-1	- <i>artham</i> > - <i>artho</i>	1 CE
		BLS4-1	- <i>artha-</i> > - <i>artha-</i>	
		BLS9-1	<i>kāraṇārthikāḥ</i> > <i>karaṇa[r]thia</i>	
		InsK-M	<i>artha-</i> > - <i>artha-</i>	2 CE
		RS-12	<i>artham</i> > <i>artha</i>	2 CE
		RS-19	<i>artha-</i> > - <i>artha-</i>	
	-thr- (1)	Aś-H	<i>artha-</i> > <i>athra-</i>	3 BCE
	-(t)th-	Aś-H	<i>artha-</i> > <i>atha-</i>	3 BCE
		BLS1-1	<i>mahārthikaḥ</i> > <i>mahathia</i>	1 CE
		BLS4-1	<i>tīrthika-</i> > <i>t[i]thiga-</i>	
		RS-14	<i>anarthikaḥ</i> > <i>aṇathio</i>	2 CE
		BLS1-1	BHS <i>prārthayi</i> > <i>prastae</i> orthographic?	1 CE
	[-st-]	BLS2-2	<i>caturthaḥ</i> > <i>caduṭho</i>	1 CE
	-(ṭ)th-	BLS4-1	- <i>arthakānām</i> > - <i>aṭheana</i>	1 CE
	-(ṭ)ṭh-			
-rd-	-rd-	InsK-M	<i>cāturdiśa-</i> > <i>cuturdiśa-</i>	3 CE

	-(d)d-	InsK-M	<i>cāturdiśa-</i> > <i>cutudiśa-</i>	1 CE
		BLS1-1	<i>cāturdiśe</i> > <i>cutudiśami</i> , <i>nirdiṣṭaḥ</i> > <i>ṇidiṭhu</i>	1 CE
		BLS4-1	<i>kardama-</i> > <i>kadama-</i>	
		BLS9-1	<i>cāturdiśaḥ-</i> > <i>cutud.[śo]-</i>	
		InsK-M	<i>cāturdiśa-</i> > <i>cutudiśa-</i>	3 CE
-rdh-	-rdh-	InsK-M	<i>sārdham</i> > <i>sardha</i>	1 BCE
		BLS9-1	<i>sārdhaṃcārin-</i> > <i>sa[r]dhacare-</i>	1 CE
		InsK-M	<i>saṃvardhaka-</i> > <i>-saṃvardhaka-</i>	2 CE
	-(d)dh-	BLS4-1	<i>bahirdhā-</i> > <i>bahidha-</i>	1 CE
		BLS9-1	<i>sārdham</i> > <i>sadha</i>	
		RS-12	<i>sārdham</i> > <i>sadha</i>	2 CE
	-(d)ḍh-	Aś-H	<i>vardhate</i> > <i>vaḍheti</i>	
	-d-	InsK-M	<i>sārdhavicārin-</i> (?) > <i>sadaviyari-</i>	3 CE
-rp-	-rp-	BLS1-1	In <i>tarpeṣi</i>	1 CE
	-(p)p-	BLS4-1	In <i>p(a<*sa>[ta]paīda-</i> (?)	
-rbh-	-rbh-	Aś-H	<i>garbhāgāra-</i> > <i>grabhagara-</i> (2) orthographic?	3 BCE
	-(h)h-	BLS4-1	<i>caturbhiḥ</i> > <i>caduhi</i>	1 CE

Other OIA combinations that do not appear in the Gāndhārī source materials are *rkh*, *rch*, *rṭ*, *rd*, *rph* and *rb*.

Overview: The Gāndhārī reflexes of OIA -rT- clusters fall into two main groups, involving either retention of the cluster or assimilation to form a geminate stop, which is recorded in the usual way through the writing of an unweakened single intervocalic stop. Only two or three Gāndhārī examples are recorded for the OIA -rk-, -rgh-, -rc-, -rj-, -rp- and -rbh- clusters, all of which show mixed results and are difficult to interpret based on the small sample size. A larger number of the source documents record reflexes for the OIA velar -rg- cluster, of which twelve (12) examples are listed in the table, and the dental -rt- (14), -rth- (15), -rd- (6) and -rdh- (8) clusters. Of these the velar and the aspirated dental clusters show mixed reflexes, while the unaspirated dental clusters mostly record geminate reflexes. Occasional variant forms such as -Cr- appear to be orthographic, while the geminate reflexes of the dental clusters are occasionally marked as retroflex stops or once with a cauda diacritic. While both reflexes have been recorded in the Aśokan and other inscriptions, as well as in the later era manuscripts by all scribes, a slight trend away

from the traditional spelling towards the geminate reflex can be discerned in the diachronic record. Metathesis, which is a common phenomenon with liquids and obstruents, is also observed occasionally, as in *kirti-* > *kitri-* (Aś-H) for example.

Discussion: Similar to the ST type clusters (§ 8.3.3.1), the pronunciation of *rT* clusters involves a period of controlled articulation during closure to the stop. As discussed in the introductory remarks (§ 8.3.4.4), this may be realised as a non-contact continuant such as the velar [ɰ], or alveolar [ɭ], in which the closure is restricted by the shaped tongue blade raised towards the velum or the alveolar ridge. Approximants have strong internal formants which provide reliable cues to the manner and place of articulation (§ 6.1), and it is expected that the *rT* clusters would possess an inherent level of stability, similar to the ST clusters which are mostly preserved among the unaspirated coronal stops in Gāndhārī. The tendency for *-rT-* clusters to assimilate in favour of the stop is similar to the MIA and Pali developments (§ 9.2.2), although this appears to have been realised as a more gradual progression in the Gāndhārī orthography. While this development has been interpreted in terms of strength-based manner assimilation in MIA (§ 9.3.1), it can also be explained in terms of the diachronic development of Sanskrit gemination patterns and strengthening of the syllable boundary (§ 9.3.4). While the *r* diacritic may have been used to mark geminates in some instances (Salomon 2000: 90), it has also been suggested that preserved forms in the manuscripts may reflect the habitual use of historic spellings in common words such as *artha-* and *sardha-*²⁷⁶. This however suggests the retention of these forms in the MIA source dialects as well, where they are more likely to have been assimilated. It is also possible that continued use of the *rC akṣaras* over several centuries indicates the retention of these clusters in speech, or at least in some dialectal divergences. On the other hand, the continued use of historic forms could also reflect spelling alternations which were retained from early stages of the sound transition and over the following centuries. It is also possible some of the retained clusters in the third century CE inscriptions may be restored or “Sanskritised” forms (§ 8.7). While a consensus of opinion suggests that the geminate forms represent the underlying language development, the actual development of *rT* clusters remains an area of uncertainty in Gāndhārī due to the variable orthographic record, which is confused to some degree by cross dialect borrowings.

²⁷⁶ Baums 2009: 162-163; Silverlock 2015: 261-262.

8.3.4.4.2 *r* + nasal clusters

Examples: Gāndhārī reflexes of OIA *r* + nasal clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.4.4.2a.

Table 8.3.4.4.2a: Gāndhārī reflexes of *r* + nasal clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>-rṇ-</i>	<i>[-rṇ-]</i>	BLS1-1	<i>svarṇa-</i> > (*śpa)[r](*ṇa)- uncertain	1 CE
	<i>-rṇ-</i>	BLS2-2	<i>uttīrṇaḥ</i> > <i>utrino</i> (This scribe writes <i>n</i> for <i>ṇ</i>)	
	<i>-(ṇ)ṇ-</i>	BLS1-1	<i>karṇa-</i> > <i>kaṇa-</i> , <i>jīrnāḥ</i> > <i>jīṇa</i>	1 CE
		BLS1-3	<i>jīrṇam</i> > <i>jīṇam</i> , <i>tīrṇaḥ</i> > <i>-tiṇo</i>	
		BLS4-1	<i>caturṇa-</i> > <i>caduṇa-</i> , <i>uttīrṇa-</i> > <i>utiṇa-</i>	
		BLS9-1	<i>avaśīrṇa-</i> > [o]śīṇa-	
		BCS5-2	<i>paripūrṇa-</i> > <i>paripuṇa-</i>	2 CE
		RS-5	<i>caturṇām</i> > <i>caḍoṇa</i>	2 CE
		RS-20	<i>caturṇām</i> > <i>caḍoṇa</i>	
		Aś-H	<i>Tāmraparṇī</i> > <i>Tambapamṇi</i> (#)	3 BCE
<i>-rm-</i>	<i>-rm-</i>	Aś-H	<i>karman-</i> > <i>krama-</i> , <i>dharmā-</i> > <i>dhrama-</i> (#)	3 BCE
		InsK-M	in <i>dharmarava-</i>	1 BCE
		BLS1-1	<i>karma(n)-</i> > <i>-karma-</i> , <i>-dharmā-</i> > <i>-dharma-</i>	1 CE
		BLS1-2	<i>dharmā-</i> > <i>dharma-</i>	
		BLS2-1	<i>-dharmāḥ</i> > <i>-dharma</i>	
		BLS2-2	<i>dharmā-</i> > <i>dharmā-</i> , <i>karma-</i> > <i>ka[r]ma-</i>	
		InsK-M	<i>dharmā-</i> > <i>dharmā-</i>	2 CE
		RS-5	<i>anudharmāḥ</i> > <i>aṇudharma</i>	2 CE
		RS-12	<i>-dharmāt</i> > <i>-dharma</i>	
		RS-14	<i>karman-</i> > <i>karma-</i> , <i>dharmā-</i> > <i>dharmā-</i>	
		RS-20	<i>dharmam</i> > <i>dharmā</i>	
		InsK-M	<i>dharmā-</i> > <i>dharmā-</i>	3 CE
	<i>-rum-</i>	InsK-M	<i>Buddhavarman-</i> > <i>Budhavaruma-</i> (#)	3 CE
	<i>-(m)m-</i>	InsK-M	<i>navakarmika-</i> > <i>navakamika-</i>	1 BCE
			<i>dharmā-</i> > <i>dhama-</i>	1 CE
		BLS1-1	<i>saddharmāḥ</i> > <i>sadhamu</i> (once only)	1 CE

		BLS1-2	<i>dharmā-</i> > <i>dhama-</i> , * <i>daurmana-</i> > <i>domaṇa-</i>	
		BLS4-1	<i>karman-</i> > <i>kama-</i> , <i>dharmā-</i> > <i>dhama-</i>	
		BLS9-1	<i>nirmohaḥ</i> > <i>ṇimoho</i>	
		RS-5	<i>karmaniyam</i> > <i>kamaṇio</i>	2 CE
		RS-12	<i>-dharmāt</i> > <i>-dhame</i> , <i>nirmāṇa-</i> > <i>ṇimaṇa-</i>	
		RS-19	<i>-karmāntaḥ</i> > <i>-kamato</i>	
		RS-20	<i>daurmana-</i> > <i>-domaṇa-</i>	
		InsK-M	<i>dharmā-</i> > <i>dhama-</i>	3 CE

OIA combinations that do not appear in the Gāndhārī source materials are *rñ* and *rn*, which are rare in OIA (Kobayashi 1999: 186). The OIA *rn* cluster mostly develops in combinations of the preverb *dus-* and verbs beginning with *n-*, forming *-rn-* through external sandhi. This differs from the internal sandhi development of *-rñ-* seen in the above examples.

Overview: Similar to the *-rT-* clusters, the Gāndhārī reflexes of OIA *-rN-* clusters fall into two main groups, indicating either retention of the cluster or assimilation to form a geminate nasal. The retroflex *-rñ-* cluster mostly assimilates as the geminate *-(ṇ)ṇ-* cluster, although the original cluster appears to be retained in two of the listed cases. The geminate is typically recorded by the writing of a single nasal, although it is marked using *anusvāra* in *Tambapaṇṇi* from the Aśokan inscriptions. Reflexes of *-rm-* either retain the original cluster or assimilate as the geminate *-(m)m-*, which is recorded as a single nasal.

Discussion: The tendency for *-rN-* clusters to assimilate in favour of the nasal is similar to the MIA developments (§ 9.2.2), although this appears to have been realised as a more gradual progression in the Gāndhārī orthography, especially in the case of the *-rm-* clusters. While this development has been interpreted in terms of strength-based manner assimilation in MIA and Pali (§ 9.3.1), it can also be explained in terms of the diachronic development of Sanskrit gemination patterns and strengthening of the syllable boundary (§ 9.3.4). Although the original cluster appears to be retained in two of the listed cases, the occasional use of pre-consonantal *r* as a marker of gemination suggests these may also be intended as geminate reflexes (Lenz 2010: 32), while the use of *anusvāra* in *Tambapaṇṇi* (Aś-H) confirms the geminate reflex. The orthographic retroflex geminate *-(ṇ)ṇ-* was most likely pronounced as dental *-n(n)-* due to the loss of retroflex conditioning by *r* and general orthographic merger of *-ṇ-* as *-n-* in Gāndhārī (§ 8.2.2.2.2). The apparent

preservation of *-rm-* reflexes occurs almost exclusively in the writing of the common Buddhist terms *karman-* and *dharmā-*, and appears to represent the acceptance of historic spellings and possible pronunciation for these terms. The alternate reflex as a geminate cluster occurs in some instances for these words, as well as in all of the other listed examples, and it seems most likely this advanced form represents the underlying language development²⁷⁷. The writing of Skt. *dharmā-* > G *dhrama-* in the Aśokan inscriptions may either be an orthographic error or a case of metathesis, while the atypical epenthetic reflex in *Buddhavarman-* > *Buddhavaruma-* appears to be a borrowing (Baums 2009: 162).

8.3.4.4.3 *r* + fricative clusters

Examples: Gāndhārī reflexes of OIA *r* + fricative clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.4.4.3a.

Table 8.3.4.4.3a: Gāndhārī reflexes of *r* + fricative clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
<i>-rś-</i>	<i>-rś-</i>	Aś-H	<i>darśana-</i> > <i>draśana-</i> (#)	3 BCE
		Aś-M	<i>-darśin-</i> > <i>-drasi-</i> (#)	
		BLS1-1	<i>darśani-</i> > <i>darśani-</i>	1 CE
		BLS2-2	<i>vidarśayati</i> > <i>vidarśayadi</i>	
		BCS5-1	<i>darśana-</i> > <i>[d]arśana-</i>	2 CE
	<i>-(ś)ś-</i>	Aś-H	<i>darśana-</i> > <i>daśana-</i>	3 BCE
		BLS1-2	<i>-darśāvā</i> > <i>-[daśa]vi</i> ; uncertain	1 CE
		BLS4-1	<i>darśana-</i> > <i>daśana-</i>	
		RS-12	<i>adarśat</i> > <i>adaśaḍi</i>	2 CE
	<i>-(ṣ)ṣ-</i>	BLS1-1	* <i>sparśeyam</i> > <i>phaṣea</i>	1 CE
		BLS1-2	* <i>sparśitvā</i> > <i>phuṣita</i> ; interpretation	
		BLS4-1	<i>sparśita-</i> > <i>phaṣida-</i>	
		BLS9-1	<i>sparśayet</i> > <i>phaṣae</i>	
		InsK-M	<i>sparśa-</i> > <i>phaṣa-</i>	2 CE
	<i>[-ṣ-]</i>	RS-20	<i>saṃparśa</i> > <i>-sapaṣa-</i>	2 CE
		RS-12	BHS <i>sparśavihāra-</i> > <i>phaṣavihara-</i>	2 CE

²⁷⁷ Baums 2009: 162-163; Silverlock 2015: 261-262.

<i>-rṣ-</i>	<i>-rṣ-</i>	Aś-H	<i>*pārṣamda > praṣamda (#)</i>	3 BCE
		BLS1-1	<i>varṣakoti- > barṣakudi-, subarṣī > suharṣi</i>	1 CE
		BLS2-1	<i>varṣavāsaḥ > varṣavaṣo</i>	
	<i>-(ṣ)ṣ-</i>	Aś-H	<i>*pārṣamda > paṣamda</i>	3 BCE
		BLS1-1	<i>vyākārṣīt > viaghāṣe</i>	1 CE
		BLS2-2	<i>varṣāvāsaḥ > vaṣavaṣo</i>	
		BLS4-1	<i>caturṣu- > caduṣu-</i>	
	<i>-(ṣ)ṣ-</i>	RS-5	<i>parikaṣṣati > parikaṣṣaḍi</i>	2 CE
<i>-rh-</i>	<i>-rah-</i>	Aś-H	<i>garha- > garaha-</i>	3 BCE
		InsK-M	<i>arhat- > arahata-</i>	1 CE
		BLS1-1	<i>arhataḥ > arahatasa</i>	1 CE
		BLS2-2	<i>antarhitah > atarahide, arhan > [ra]hado</i>	
		BLS4-1	<i>etarhi > edarahi</i>	
	<i>-r-</i>	Aś-H	<i>garhana- > garana-</i>	3 BCE
	<i>-h-</i>	BCS5-2	<i>tarhi > tahi, yarhi > yahi</i> (indeclinables)	

One other OIA combination which does not appear in the Gāndhārī source materials is *rs*. Similar to the *rn* cluster this is uncommon in OIA (Kobayashi 1999: 186), compared with the normal sandhi development of *-rṣ-* seen in the tabulated examples²⁷⁸.

Overview: Similar to the *-rT-* and *-rN-* clusters, reflexes of OIA *-rś* and *-rṣ-* clusters fall into two main groups. These indicate retention of the cluster or assimilation to form a geminate sibilant in the same place of articulation, although the retroflex geminate *-(ṣ)ṣ-* is also recorded as an alternate development for the *-rś-* clusters. The geminates are recorded as single intervocalic sibilants. Gāndhārī reflexes of the OIA glottal fricative *-rh-* typically involve insertion of an epenthetic vowel as *-rah-*, although in the case of Skt. *arhat-* > G *arahata-* this is likely to be a borrowing or retention of a venerable form.

Discussion: Compared with the *-rT-* and *-rN-* clusters, which involve full oral closure of the following obstruent, realisation of *-rS-* clusters requires an adjustment of the tongue between different positions of partial closure for the two continuants. The development of geminate sibilants suggests a regressive assimilation of manner in which the tongue anticipates the closure of the following fricative. The *-rS-* clusters were generally

²⁷⁸ Dental *rs* typically remains in some reduplicated stems of the 3rd class verbal root \sqrt{sr} , as for example in *sisrate* (present indicative, 3rd person, plural [MW s.v.]), among other derivatives.

assimilated in favour of the fricative in MIA (§9.2.2), although this appears to have developed only partially in Gāndhārī where a mixed pattern of stable and assimilated reflexes is seen. This development has been interpreted in terms of strength-based manner assimilation in MIA and Pali (§ 9.3.1), and is also be explained in terms of the diachronic development of Sanskrit gemination and strengthening of the syllable boundary (§ 9.3.4). The OIA *-rś-* cluster is typically recorded in various forms of *-darśa-* and *-sparśa-*, which are derivatives of the Sanskrit verbal roots $\sqrt{drś}$ and $\sqrt{sprś}$ respectively, involving strengthening of vocalic *-ṛ-* to *-ar-*. While both retained *-rś-* and geminate *-(ś)ś-* reflexes are recorded for *-darśa-*, the typical development of *-sparśa-* involves the geminate *-(ś)ś-* reflex only. This is recorded in several manuscripts, and it is possible this represents the reflex of an allophonic pronunciation of *-rś-* conditioned as *-rṣ-*, while the writing of *-ṣ-* may be a way of recording the geminate. The writing of *darśana* > *draśana-*, *-darśin-* > *-drasi-* and *pārṣamda-* > *praṣamda-*²⁷⁹, along with *kirti-* > *kitri-* in the *r* + stop clusters *karma-* > *krama-* and *dharma-* > *dhrama-* in the *r* + nasal clusters, and *sarva-* > *savra-* and *pūrva-* > *pruva-* in the *r* + semi-vowel clusters (Table 8.3.4.4a), suggests a persistent recording of metathesis in the Aśokan inscriptions.

The assimilation of these clusters in the other MIA languages suggests that conservative forms are less likely to be borrowings, while the continued use of both reflexes over several centuries may reflect spelling alternations which developed during the early stages of the sound transition. Like the recording of variable reflexes for the *Sr* and *rT* clusters (§§ 8.3.3.4.4, 8.3.4.4.1), it is possible that both forms were subsequently retained and used freely in the diachronic record or in dialectical divergences. It is also possible the apparent retention of stable forms may relate to the occasional use of pre-consonantal *r* as a marker of gemination (Lenz 2010: 32). The epenthetic development of *-rh-* > *rah-* may reflect the recording of an allophonic realisation of the schwa vowel in the articulatory space between the two consonants in the cluster (§ 10.1.1), although as noted previously Skt. *arhat-* > G *arahata-* is likely to be a borrowing or retention of a venerable form.

8.3.4.4.4 *r* + semi-vowel clusters

Examples: Gāndhārī reflexes of OIA *r* + semi-vowel clusters recorded in the source materials, along with a selection of typical examples, are presented in Table 8.3.4.4a.

²⁷⁹ Hultzsch 1925: lxxxix, 188, 198; Mehendale 1948: 25.

Table 8.3.4.4a: Gāndhārī reflexes of *r* + semi-vowel clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 * = unattested equivalent (#) atypical reflexes		Era
-rl-	-(l)l-	BLS9-1	<i>nirlolupaḥ</i> > <i>ṇiloluo</i> , <i>durlabhāḥ</i> > <i>dulabha</i>	1 CE
-rv-	-rv-	InsK-M	<i>sarva-</i> > <i>sarva-</i>	1 BCE
			<i>sarva-</i> > <i>sarva-</i>	1 CE
		BLS1-1	<i>sarva-</i> > <i>sarva-</i>	
		BLS1-2	<i>gandharva-</i> > <i>ghadharva-</i> , <i>sarva-</i> > <i>sarva-</i>	1 CE
		BLS2-1	<i>sarvam</i> > <i>sarvo</i> / <i>sarva</i>	
		BLS2-2	<i>sarvam</i> > <i>sarva</i> ; in several forms	
		BLS9-1	<i>sarva-</i> > <i>sarva-</i>	
		InsK-M	<i>sarva-</i> > <i>sarva-</i> , <i>pūrva-</i> > <i>purva-</i>	2 CE
		RS-5	<i>sarva-</i> > <i>sarva-</i>	2 CE
		RS-12	<i>sarvaśaḥ</i> > <i>sarvaśa</i>	
		RS-14	<i>pūrva-</i> > <i>parva-</i>	
		RS-20	<i>sarve</i> > <i>sarva</i>	
		BCS5-1	<i>rv</i> > <i>rv</i> , <i>rv</i> > <i>vr</i> (#)	2 CE
		InsK-M	<i>sarva-</i> > <i>sarva-</i>	3 CE
	-vr-	Aś-H	<i>sarva-</i> > <i>savra-</i> (#)	3 BCE
		RS-5	<i>nirvidā</i> > <i>ṇivriḍa</i> (#)	2 CE
	-ruv-	Aś-H	<i>pūrva-</i> > <i>pruva-</i> (#)	3 BCE
		Aś-H	<i>nirvarta-</i> > <i>nivata-</i>	
	-(v)v-	InsK-M	<i>nirvāṇa-</i> > <i>ṇivaṇa-</i>	1 CE
		BLS1-2	<i>sarva-</i> > <i>sava-</i>	1 CE
		BLS4-1	<i>nirvāṇa-</i> > <i>ṇivaṇa-</i> , <i>sarva-</i> > <i>sava-</i>	
		RS-19	<i>nirvāṇa-</i> > <i>ṇivaṇa-</i>	2 CE
		RS-20	<i>sarve</i> > <i>[s(*a)v(*a)]</i>	
		InsK-M	<i>sarvatratā-</i> > <i>savatrata-</i>	2 CE
	-ṃv-	InsK-M	<i>sarva-</i> > <i>saṃva-</i> (#); assume marks geminate	2 CE
		RS-14	<i>sarvottamaḥ</i> > <i>sa[b]utamo</i> (?)	2 CE
	-(b)b- (?)			
-ry-	-ry-	InsK-M	<i>ācārya-</i> > <i>acarya-</i> , <i>bhāryā-</i> > <i>bharya-</i>	1 CE
			<i>ācārya-</i> > <i>acarya-</i> , <i>bhāryā-</i> > <i>bharya-</i>	2 CE
			<i>ācārya-</i> > <i>acarya-</i>	3 CE

	<i>-riy-</i>	BLS2-1	<i>kāryam</i> > <i>karya</i>	1 CE
		BLS2-2	<i>kāryam</i> > <i>karyam</i>	
		BLS9-1	<i>-vīrya-</i> > <i>-vīrya-</i>	
		BCS9-1	<i>kārya-</i> > <i>karya-</i>	2 CE
		Aś-H	<i>ānamtaryena</i> > <i>anamtariyeṇa</i>	3 BCE
		Aś-M	<i>mādhurya-</i> > <i>madhuriya-</i>	3 BCE
	<i>-ri(y)-</i>	InsK-M	<i>ācārya-</i> > <i>ayaria-</i> <i>-ry-</i> > <i>*riy-</i> > <i>-ri-</i>	1 CE
		BLS1-1	<i>vīryavatām</i> > <i>viri[v](*ata)ṇa</i> ; uncertain	1 CE
		BLS1-2	<i>vīrya-</i> > <i>(*vi)ria-</i>	
		BLS2-2	<i>vīryendriyaḥ</i> > <i>viridrigo</i>	
		BCS5-1	<i>jāgaryā</i> > <i>jagaria</i>	2 CE
		InsK-M	<i>bhāryā-</i> > <i>bharia-</i>	3 CE
	<i>-rig-</i> <i>-(y)y-</i>	BLS2-2	<i>sūryam</i> > <i>surigo</i> (#)	1 CE
		Aś-H	<i>-āryeṣu</i> > <i>-ayeṣu</i>	3 BCE
		Aś-M	<i>marya-</i> > <i>maya-</i>	
		InsK-M	<i>bhāryā-</i> > <i>-bhaya-</i>	1 BCE
		BLS4-1	<i>paryeṣati</i> > <i>payeṣadi</i>	1 CE
		RS-5	<i>paryantam</i> > <i>payata</i>	2 CE
		BCS5-1	<i>karye</i> > <i>kaye</i>	2 CE

All OIA combinations appear in the Gāndhārī source materials.

Overview: Similar to the *-rT-* and *-rS-* clusters, the Gāndhārī reflexes of OIA *r* + semi-vowel clusters show mixed results, indicating both retention and assimilation of the clusters in favour of the second consonant. The majority of recorded cases are for reflexes of the OIA *-rv-* and *-ry-* clusters, in which the original cluster appears to have been retained more often than assimilated. Assimilated clusters such as *rv* > *(v)v* and *ry* > *(y)y* are represented by single *akṣaras* in the Kharoṣṭhī script, and are assumed to represent geminates based on MIA parallels. There are some atypical reflexes which appear to involve metathesis or the insertion of epenthetic vowels, such as *-rv-* > *-vr-* or *-ruv-*, and *ry* > *-riy-* or *-ri(y)-* for example. There are only two examples involving the *-rl-* cluster which have both been assimilated to form the assumed *-(l)l-* geminate.

Discussion: All of the semi-vowels have strong internal formant structures, and in this context, there appears to be little to distinguish between them on the basis of internal

acoustic signals, making them more reliant on the formant transitions to adjacent vowels for their comparative perceptual identity (§ 6.1). Transitions of the second consonant to the following vowel appear to offer the most salient cues to the place of articulation of consonants. Since the semi-vowels are typically continuants with different places of articulation, it is possible that the tendency to regressive place assimilation in these clusters, as in *-rv-* > *-(v)v-* and *-ry-* > *-(y)y-*, reflects the stronger perceptual cues of the second consonant. This approach does not however explain the progressive assimilation of the *-vy-* and *-vr-* clusters seen in some of the Gāndhārī examples of *vR* clusters (§ 8.3.4.2.4), which have been explained in terms of the relative weakness of *y* and *r* in the semi-vowel strength hierarchy (§ 6.1.3), and the *-rv-* > *-(v)v-* and *-ry-* > *-(y)y-* geminate reflexes are also consistent with this approach. These reflexes are also similar to the typical MIA developments of *-vv-* / *-bb-* and *-yy-* / *-jj-* respectively (§ 9.2.2). At the same time the retention of *-rv-* and *-ry-* in many of the Gāndhārī reflexes suggests that both clusters possess an inherent perceptual stability. Preserved forms are mostly seen in common words such as *sarva-* and *ācārya-*, and may reflect retained historical forms for these words (Baums 2009: 183), although both conservative and advanced forms are sometimes seen in the same manuscript, such as *sarva-* and *sava-* in BLS1-2 and RS-20 for example. In light of the progressed MIA developments, it is also possible that the written geminates were borrowings, or as suggested previously, both forms may reflect alternate spellings that were retained from an earlier transitional stage in the language. The relatively common insertion of epenthetic *i* in the development of *-vy-* > *-vi(y)-* may also be interpreted as an indication that the cluster was retained. While an epenthetic development such as *-ry-* > *-riy-* > *-ri(y)-* > *-ri-* is possible²⁸⁰, a direct allophonic recording of *i* for *y* is also possible due to their shared articulatory and acoustic properties (§ 8.3.1.4.3), and in this context the original pronunciation appears to have been retained.

While there is a tendency for the OIA *r* + semi-vowel clusters to assimilate in Gāndhārī according to the MIA strength hierarchy of *l* > *v* > *y* > *r*, there are sufficient counter examples where the original cluster is retained to suggest that this type of assimilation may not have been fully progressed in Gāndhārī. And while the retained clusters may be explained in terms of historic forms, the recording of *-vy-* as *-vi-* is suggestive of a retained pronunciation, while the writing of advanced forms may also be interpreted as borrowings from other MIA languages. The general impression is that both

²⁸⁰ Mehendale 1948: 308; Salomon 2008: 123.

the *-rv-* and *-ry-* clusters are typically retained, while transitional developments towards geminate clusters, whether through local language change or borrowed forms, are evident in the phonological record.

8.3.4.4.5 Summary of *r* + C cluster developments

A summary of the Gāndhārī reflexes of OIA *r* + C clusters is presented in Table 8.3.4.4.5.

Table 8.3.4.4.5: Summary of *r* + C cluster developments in Gāndhārī

OIA Cluster	Suggested G phoneme	Assumed sound	Alternate representations of Gāndhārī phonemes These are classified as historic forms (h), transitional allophones (tr), advanced forms (adv), alternate allographs (alt), or borrowed forms (borr). (Bracketed items) are irregular or infrequent.
<i>/-rT-/</i> T = <i>g, t, th, d, dh</i>	<i>/-TT-/</i> (<i>/-rT-/</i> alt. mixed reflexes)	[-TT-]	<p>The OIA <i>-r{g, t, th, d, dh}</i>- clusters are most widely represented in the Gāndhārī reflexes. Although the orthographic record is at times confused, the following typical pattern is suggested:</p> <p><i>/-TT-/</i> \equiv { <i>/-T-/</i>^{adv}, <i>/-rT-/</i>^h }</p> <p>The geminates are typically represented by a single stop, while the <i>/-rT-/</i>^h reflexes appear to be historic forms in many cases, although the record is mixed. Atypical reflexes also occur in some cases:</p> <p><i>/-g-/</i>^{adv} \equiv { <i>/-g-/</i>, (<i>/-g-/</i>, <i>/-gh-/</i>) }</p> <p><i>/-rg-/</i>^h \equiv { <i>/-rg-/</i>, (<i>/-rVg-/</i>, <i>/-gr-/</i>) }</p> <p><i>/-t-/</i>^{adv} \equiv { <i>/-t-/</i>, (<i>/-t-/</i>, <i>/-tʰ-/</i>) },</p> <p><i>/-rt-/</i>^h \equiv { <i>/-rt-/</i>, (<i>/-tr-/</i>) }</p> <p><i>/-th-/</i>^{adv} \equiv { <i>/-th-/</i>, (<i>/-tʰ-/</i>, <i>/-tʰh-/</i>) },</p> <p><i>/-rth-/</i>^h \equiv { <i>/-rth-/</i>, (<i>/-thr-/</i>) }</p> <p><i>/-dh-/</i>^{adv} \equiv { <i>/-dh-/</i>, (<i>/-dʰ-/</i>, <i>/-d-/</i>) }</p> <p><i>/-rdh-/</i>^h \equiv <i>/-rdh-/</i></p> <p>Reflexes of OIA (<i>-r{k, gh, c, j, p, bh}</i>-) clusters are only recorded in a small number of cases, and reflexes include both assimilation and retention:</p> <p>(<i>/-TT-/</i>) \equiv ({ <i>/-T-/</i>^{adv}, <i>/-rT-/</i>^h })</p>
<i>/-rT-/</i> T = <i>k, gh, c, j, p, bh</i>	<i>/-TT-/</i> (<i>/-rT-/</i> alt)		

<i>/-rṇ-/</i>	<i>/-ṇṇ-/</i>	[-ṇṇ-]	<p><i>/-ṇṇ-/</i> $\equiv \{ /-ṇ-/^{adv}, (/rṇ-/^{alt}, /mṇ-/^{alt}) \}$</p> <p>The geminate is typically represented by a single nasal. Occasional use of the <i>r</i> diacritic is interpreted as a marker of the geminate, while the use of <i>anusvāra</i> follows the common inscriptional form.</p>
<i>/-rm-/</i>	<i>/-mm-/</i>	[-mm-]	<p><i>/-mm-/</i> $\equiv \{ /-m-/^{adv}, /rm-/^h, (/rum-/^{alt}) \}$</p> <p>The geminate is typically represented by a single nasal, while preserved forms appear to be historic spellings of common terms. The insertion of epenthetic <i>u</i> appears to be a borrowing.</p>
<i>/-rś-/</i>	<i>/-śś-/</i> (<i>/-rś-/</i> alt. mixed)	[-çç-]	<p><i>/-śś-/</i> $\equiv \{ /-ś-/^{adv}, /rś-/^h, /-ṣ-/^{alt} \}$</p> <p>The geminate is typically represented by a single sibilant. Alternate <i>/-rś-/</i> reflexes appear to be historic, while retroflex <i>ṣ</i> is possibly allophonic.</p>
<i>/-rṣ-/</i>	<i>/-ṣṣ-/</i> (<i>/-rṣ-/</i> alt. mixed)	[-ṣṣ-]	<p><i>/-ṣṣ-/</i> $\equiv \{ /-ṣ-/^{adv}, /rṣ-/^h, (/ṣ-/^{alt}) \}$</p> <p>The geminate is typically represented by a single sibilant, while <i>/-rṣ-/</i>^h reflexes appear to be historic forms and diacritic <i>ṣ</i> may record the geminate.</p>
<i>/-rh-/</i>	<i>/-rh-/</i>	[-ɹəfi-]	<p><i>/-rh-/</i> $\equiv \{ /-rah-/^h, (/r-/^{alt}) \}$</p> <p>It is assumed that the cluster is retained in speech and that the epenthetic <i>/-rah-/</i>^h reflex records allophonic pronunciation, while the allograph <i>/-r-/</i>^{alt} is an orthographic error.</p>
<i>/-rl-/</i>	<i>/-ll-/</i>	[-ll-]	<p><i>/-ll-/</i> $\equiv /-l-/^{adv}$</p> <p>The geminate is typically represented by a single consonant. This case is based on a very small sample.</p>
<i>/-rv-/</i>	<i>/-rv-/</i> (<i>/-vv-/</i> alt. mixed)	[-ɹv-]	<p><i>/-rv-/</i> $\equiv \{ /-rv-/^h, /-v-/^{adv}, (/vr-/^{alt}) \}$</p> <p>While the <i>/-rv-/</i>^h reflexes are historic, it remains uncertain whether <i>/-v-/</i>^{adv} are advanced forms or borrowings, and the overall status of the development is uncertain. <i>/-vr-/</i>^{alt} appears to be an orthographic error.</p>
<i>/-ry-/</i>	<i>/-ry-/</i> (<i>/-yy-/</i> alt. mixed)	[-ɹj-]	<p><i>/-ry-/</i> $\equiv \{ /-ry-/^h, /-riy-/^{alt}, (/y-/^{adv}) \}$</p> <p>Due to the mixed evidence, it has been assumed, possibly conservatively, that the cluster is retained and the epenthetic reflex <i>/-riy-/</i>^{tr} records allophonic pronunciation, while the <i>/-y-/</i>^{adv} represents advanced spellings which are possibly borrowings.</p>

8.3.5 Three consonant clusters

Gāndhārī reflexes of OIA three consonant clusters recorded in the source materials, and which have clearly attested Sanskrit equivalents, are presented in table 8.3.5. These have been organised according the manners of the first and second elements, following the ordering sequence used in the previous sections describing the two consonant clusters. Unless noted otherwise all clusters occur in word medial positions.

Table 8.3.5: Gāndhārī reflexes of OIA three consonant medial clusters

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī		Era
For source references refer to Table 8.1 * = unattested equivalent #CCC- = word initial cluster				
Stop +				
TT+				
<i>ttv</i>	<i>tv</i>	InsK-M	<i>sattva-</i> > <i>satva-</i>	1, 2, 3 CE
		BCS5-1	<i>sattva-</i> > <i>satva-</i>	2 CE
<i>ḍdv</i>	<i>ḍv</i>	RS-20	<i>ṣaḍ-dvāra-</i> > <i>ṣa-dvaro-</i>	2 CE
<i>bdhv</i>	<i>dhv, dh</i>	BLS4-1	<i>labdhvā</i> > <i>ladhva, ladha</i>	
TN+				
<i>tmy</i>	<i>t(t)v</i>	BLS1-1	<i>māhātmya-</i> > <i>mahatvo-</i>	1 CE
TS+				
<i>kṣm</i>	<i>hom</i>	BLS4-1	<i>sūkṣma-</i> > <i>suhoma-</i>	1 CE
<i>kṣy</i>	<i>kṣ</i>	BLS1-1	<i>mokṣyatha</i> > <i>mukṣasa</i>	
		BLS1-2	<i>*drakṣyantu</i> > <i>dhrekṣatu</i>	
<i>kṣv</i>	<i>ṣm</i>	InsK	<i>ikṣvāku-</i> > <i>iṣmahu-</i> (Salomon and Baums 2007)	1 CE
<i>tsth</i>	<i>tth</i>	Aś-H	<i>ut-stha-</i> > <i>uthana-</i>	3 BCE
TR+		none recorded		
Nasal +				
NT+				
<i>ṁkṣ</i>	<i>kṣ</i>	BLS4-1	<i>kaṅkṣa-</i> > <i>kakṣa-</i> , <i>saṁkṣepa-</i> > <i>sakṣeva-</i>	1 CE
		BCS5-1	BHS <i>saṁkṣipta-</i> > <i>sakṣita-</i>	2 CE
<i>ṅkṣ</i>	<i>(ṁ)kṣ</i>	RS-12	<i>ākāṅkṣa-</i> > <i>agakṣa-</i>	2 CE
<i>ṁkl</i>	<i>kil</i>	BLS4-1	<i>saṁkliṣya-</i> > <i>sakileśa-</i>	1 CE
<i>ṁkhy</i>	<i>kh</i>	BLS4-1	<i>saṁkhyā-</i> > <i>sakha-</i>	1 CE
	<i>ṁkh</i>	BCS5-1	<i>asaṁkhyāta-</i> > <i>asaṁkhada-</i>	2 CE

<i>mgr</i>	<i>gr</i>	BLS4-1	<i>saṃgraha-</i> > <i>sagraha-</i>	1 CE
<i>mjñ</i>	<i>ñ(ñ)</i>	RS-12	<i>-saṃjña-</i> > <i>-saña-</i>	2 CE
	<i>ñ</i>	BLS4-1	<i>saṃjña-</i> > <i>saña-</i>	1 CE
<i>mtv</i>	<i>tv</i>	BLS1-1	<i>saṃtvaritaḥ</i> > <i>satvaridu</i>	1 CE
<i>ntr</i>	<i>tr</i>	BLS1-2	<i>āmantrayate</i> > <i>amatredi</i>	1 CE
		BLS2-2	<i>nimantrita-</i> > <i>nimatrido-</i> , <i>mantrita-</i> > <i>matr[ido-]</i>	
		BLS4-1	<i>*mantrayate</i> > <i>matredi</i>	
		RS-5	<i>antra-</i> > <i>atra-</i>	2 CE
<i>ndy</i>	<i>nd / ndiy</i>	As-H	<i>Pāṇḍya-</i> > <i>Paṁda-</i> / <i>Paṁdiya-</i> (#)	3 BCE
<i>ndr</i>	<i>dhr</i>	BLS1-2	<i>indriya-</i> > <i>-idhri-</i>	1 CE
	<i>dr</i>	BLS2-2	<i>indra-</i> > <i>idra</i> , <i>vīryendriya-</i> > <i>viridrigo-</i>	
		BLS4-1	<i>indriya-</i> > <i>idria-</i>	
		RS-5	<i>indriya-</i> > <i>hydria-</i>	2 CE
<i>mpr</i>	<i>(m)pr</i>	RS-12	BHS <i>saṃpraharṣa-</i> > <i>saprahasa-</i>	2 CE
	<i>pr</i>	BLS4-1	<i>asaṃprakhyāna-</i> > <i>asapraḥka(*ṇa)-</i>	1 CE
NS+				
<i>msk</i>	<i>kh</i>	BLS4-1	<i>saṃskara-</i> > <i>sakhara-</i>	1 CE
<i>msth</i>	<i>ḥh</i>	BLS4-1	<i>-saṃsthāna-</i> > <i>-saḥhana-</i>	
<i>msth</i>	<i>mth</i>	BCS5-2	<i>saṃsthita-</i> > <i>saṃthida-</i>	2 CE
Fricative +				
ST +				
<i>skr</i>	<i>rkh</i>	BLS1-1	<i>*abhiniṣkram</i> > <i>abhiṇirkham</i>	1 CE
	<i>kr</i>	RS-20	<i>niṣkramanti</i> > <i>ṇakramati</i>	2 CE
<i>stv</i>	<i>sp</i>	RS-5	<i>dr̥ṣṭvā</i> > <i>drispa</i>	2 CE
		RS-12	<i>dr̥ṣṭvā</i> > <i>drispa</i>	
<i>stv</i>	<i>st</i>	BLS1-2	<i>*daurmanastva-</i> > <i>domaṇastu-</i>	1 CE
<i>sty</i>	<i>st</i>	RS-5	<i>śrāvastya-</i> > <i>śavasti-</i>	2 CE
<i>str</i>	<i>str</i>	BLS2-2	<i>sastratega-</i> > <i>sastrad(*ega-)</i>	1 CE
<i>#str-</i>	<i>#istr-</i>	BLS1-1	<i>strikuṇapa-</i> > <i>istrikuṇavu-`</i>	1 CE
		BLS4-1	<i>strīsaṃjña-</i> > <i>istrisaña-</i>	1 CE
<i>#spr-</i>	<i>#pr-</i>	BLS1-2	BHS <i>spraṣṭavya-</i> > <i>proṭhabu-</i>	1 CE
SN +			none recorded	
SR+			none recorded	
RT+				

<i>rtm</i>	<i>t̥</i>	BLS4-1	<i>vartman-</i> > <i>vaṭa-</i>	1 CE
<i>rty</i>	<i>c</i>	BLS4-1	<i>martyaka-</i> > <i>macaga-</i>	
<i>rthy</i>	<i>rth</i>	BCS5-1	<i>nisāmarthya-</i> > <i>ṇisanartha-</i>	2 CE
<i>rdhv</i>	<i>(d)dv</i>	RS-5	<i>ūrdhva-</i> > <i>udva-</i>	

In most cases the OIA clusters are formed by the addition of a final semi-vowel to a typical two consonant cluster, while final stops, nasals or fricatives only occur in a small number of cases, giving rise to different syllable divisions in the clusters, which are discussed further below²⁸¹. The following patterns of assimilation are observed.

8.3.5.1 Stop + clusters

The only reflexes of OIA TT+ clusters involve a following *v*, as in TT*v*, which are typically preserved in the form Tv, where the stops are represented by a single *akṣara* following the normal Kharoṣṭhī practice (§ 8.3.1.1). It is also possible that the stops were assimilated to form a geminate, as in *b.dhv* > *(d).dhv*, or retained as such, as in *t.tv* > *(t).tv*, and were. These developments are similar to the OIA Tv clusters which are preserved (§ 8.3.1.4.2). In all cases the OIA syllable division occurs between the two stops, and the Tv onset to the second syllable is preserved.

Development of the single TMR cluster, *t.my* > *t.(t)v*, possibly follows the pattern of loss of nasalisation and deocclusion of the labial *m*, similar to *atma* > *atva* (§ 8.3.1.2), followed by loss or assimilation of the *y* (§ 8.3.4.2.4). Concurrent strengthening of the *t* as a geminate across the syllable boundary may also occur in order to preserve the syllable structure, as in *t.my* > *t.vy* > *t.ṇ^y* > *t.tv*.

A number of different developments appear to occur in the TS+ clusters:

- *kṣy* > *kṣ* in *mokṣyatha* > *muḥṣasa* follows a similar pattern to the *kṣ* clusters which are preserved (§ 8.3.1.3), and the *ṣy* clusters where *y* is deleted or assimilated (§ 8.3.3.4.3). *kṣy* may develop as a unitary segment similar to *kṣ* (Appendix A4), possibly forming a syllable onset similar to the *kṣ* and *ṣy* clusters, and this structure is maintained.
- *kṣ.v* > *ṣ.m* in *ikṣvāku-* > *iṣmahu-* suggests a different syllable division, with weakening of *kṣ* > *ṣ* in the coda position, concurrent with occlusion and nasalisation of the labial *v* > *m* in the onset.

²⁸¹ Syllable divisions are assessed in terms of sonority profiles only in this analysis. Refer also to §§ 4.5, 8.4.

- *t.sth* > (*t*).*th* in *ut-stha-* > *uthana-*, where the *sth* onset component follows a similar development to the two consonant cluster (§ 8.3.3.1), while the initial *t* in the coda may be retained but unwritten.
- The development of *kṣm* > *hom* appears to be a borrowing (Baums 2009: 184).

8.3.5.2 Nasal + clusters

In the majority of cases nasalisation precedes a two consonant cluster of the stop plus type, and apart from *m.jñ* all of these fall into the N.T{S, R} category having rising sonority in the onset syllable structure. It is assumed that the nasal coda element is preserved in the Gāndhārī reflexes, but was not recorded following the normal practice in the Kharoṣṭhī script (§ 8.3.2.1). The onset clusters typically develop in accordance with the normal two consonant cluster reflexes, as for example:

- *m.kṣ* > (*m*).*kṣ* in *kaṅkṣa-* > *ka(m)kṣa-*, where *kṣ* is preserved (§ 8.3.1.3),
- *m.kl* > (*m*).*kl* in *saṃkliśya-* > *sa(m)kileśa-*, involving insertion of an epenthetic vowel (§ 8.3.1.4.1).
- *m.kh-y* > (*m*).*kh* in *saṃkhyā-* > *sa(m)kha-*, involving the loss of *y* (§ 8.3.1.4.3).
- *m.Tr* > (*m*).*Tr* in Skt. *mantra* > G *ma(n)tra*, where *r* is retained (§ 8.3.1.4.4).
- The *m.jñ* clusters also develop in a similar way as *m.jñ* > (*m*).*ñ*, but with palatalisation of *anusvāra* as *ñ.ñ* (§ 8.3.1.2).

The remaining clusters in this group involve the development of the OIA N.ST type clusters *m.sk* > (*m*).*kh*, *m.sth* > (*m*).*th* and *msth* > *mth*, where the ST pairs form onsets despite their falling sonority. In all cases the ST onset develops in similar ways to the two consonant clusters (§ 8.3.3.1), while the assumed nasal coda in the prior syllable is unwritten in some cases, similar to the N.T{S, R} clusters.

8.3.5.3 Fricative + clusters

All of the OIA clusters in this group are of the STR type, with two of these occurring in word initial positions. While ST clusters are able to form word initial onsets, they remain hetero-syllabic in medial positions (Appendix A4), and all of the OIA medial clusters in this group have an S.TR syllable division. In all cases but one the S coda is

preserved, and the TR onsets show variable developments, but generally similar to one of the developments recorded for the two consonant clusters, as for example:

- $\dot{s}.tv > s.p$ in $\dot{d}r\dot{s}tv\bar{a} > drispa$, where tv develops as p , which is recorded in some of the two consonant cluster reflexes (§ 8.3.1.4.2).
- $s.tv > s.t(u)$ in $*daurmanastva- > domaṇastu-$, where tv appears to develop as tu .
- $s.ty > s.t$ in $\acute{s}r\bar{a}vastyā- > \acute{s}avasti-$, where the y is elided which is consistent with the typical two cluster reflexes, although in the typical cases the retained stop is mostly palatalised as c (§ 8.3.1.4.3).
- $s.tr > s.tr$ in $sastratega- > sastrad(*ega-)$, where the final r is retained, consistent with the typical two consonant cluster reflexes (§ 8.3.1.4.4).

The word initial cases form syllable onsets, and develop in different ways:

- $strī- > is.tri-$ in $strīkuṇapa- > istrikuṇavu-$, where the cluster is retained and divided by the insertion of a word initial prosthetic vowel. Such changes are seen on a cross linguistic basis (Campbell 2016: 30), and in this case enhances or strengthens the first syllable by the addition of a new initial syllable.
- $spr- > pr-$ in BHS $spraṣṭavya- > proṭhabu-$, where the final r is preserved, which is consistent with the typical two consonant cluster reflexes (§ 8.3.1.4.4). The initial s is also lost, but without concurrent aspiration of the stop which is common in the two consonant clusters (§ 8.3.3.1) and is possibly suppressed by the following r .

8.3.5.4 Semi-vowel + clusters

All of the OIA clusters in this group are of the $rT\{m, y, v\}$ type with an initial r followed by a two consonant cluster of the TC type. Sonority falls between the r and T, and then rises to the final consonant, and by the syllable contact law (§ 4.5) the syllable division falls between r and T, giving an $r.T\{m, y, v\}$ structure. The recorded reflexes of the OIA rT clusters show a mixture of retained forms and TT geminates, the latter reflect the most advanced forms (§ 8.3.4.4.1). The development of the three consonant clusters varies, and appears to depend upon the particular scribe, as for example:

- $r.tm > (t).ṭ$ in $vartman- > va(t)ṭa-$ and $r.ty > (c).c$ in $martyaka- > macaga-$ are recorded by the BL scribe 4. In these cases, the rt clusters are resolved as geminates while the final consonant is lost. In the latter case this appears to have been absorbed through

palatalisation of the geminate stop as *c.c*, while it has been suggested that the retroflex articulation recorded in the former case is possibly a borrowed form (Baums 2009: 184).

- *r.thy* > *r.th* in *nisāmarthya-* > *ṇisanartha-* is recorded by the BC scribe 5. In this case the initial *r* is retained, while the final *y* is lost as in common in Gāndhārī, but without any further gemination or palatalisation of the stop. This may reflect a conservative form, but written without the *y* due to orthographic laxity (§ 8.3.1.4.3).
- *r.dhv* > (*d*).*dv* in *ūrdhva-* > *udva-* is recorded by the RS scribe. This example suggests a more typical development in which the initial *r* is lost, while the final *v* is retained (§ 8.3.1.4.2) and the stop develops as a geminate across the syllable boundary.

8.3.5.5 Overview

As observed by Baums (2009: 184-185), the OIA syllable division in three consonant clusters appears to play a role in how they develop in Gāndhārī. Once the OIA syllable divisions are considered, the two consonant components of the cluster tend to follow a similar development as the same stand-alone two consonant clusters. The limited number of clear examples combined with variabilities in the Kharoṣṭhī record, such as the non-recording of *anusvāra* or geminates, makes it difficult to draw clear conclusions about the patterns of development, however. The most consistent developments are seen in the N.T{S, R} and N.ST clusters, while the S.TR clusters are also reasonably consistent in their changes, with both groups retaining the original OIA syllable structure.

8.3.6 Summary of Gāndhārī reflexes of OIA consonant clusters

8.3.6.1 Word initial clusters

The Gāndhārī reflexes of OIA word initial two consonant clusters are summarised in Table 8.3.6.1. Clusters have been grouped in terms of combinations of manner, based on the MIA and Pali consonant strength hierarchies, and are recorded in summary form to represent the most typical developments. In this table blank cells are those in which Sanskrit clusters do not exist, and changes are presented in the form of Skt. cluster > G reflex. Where there are no attested Gāndhārī forms this is noted. In most cases the Sanskrit word initial clusters remained stable in Gāndhārī, and processes of lenition and deletion of the second more sonorant consonant, like those which developed in MIA (§ 9.2), had not developed to any significant extent in Gāndhārī. Changes occurred in some instances, however, which are highlighted in bold in the table and discussed further below:

Table 8.3.6.1: Typical Gāndhārī reflexes of OIA word initial two consonant clusters

	Stop (T)	Nasal (N)	Fricative (S)	<i>l</i>	<i>v</i>	<i>y</i>	<i>r</i>
Stop (T) +	None attested in Skt. or G	<i>jñ-</i> > <i>ñ-</i>	<i>kṣ-</i> > <i>-kṣ-</i>	<i>Tl-</i> > <i>Tl-</i> <i>Tv-</i> > <i>Tv-</i>		<i>Ty-</i> > <i>T-</i>⁽¹⁾ <i>Tr</i> > <i>Tr</i>	
Nasal (N) +				Skt. <i>ml-</i> none attested in G		<i>mr-</i> > <i>mr-</i>	
Fricative (S) +	<i>ST-</i> > <i>T^h-</i> <i>ST^h-</i> > <i>T^h-</i> (2)	<i>Sn-</i> > <i>Sn-</i> <i>Sm-</i> > <i>Sp-</i> (2)		<i>Sl-</i> > <i>Sl-</i> <i>Sv-</i> > <i>Sp-</i> / <i>Sv-</i>		<i>Sy-</i> > <i>Sy-</i> <i>Sr-</i> > <i>S-</i>	
<i>l</i> +							
<i>v</i> +						<i>vy-</i> > <i>vy-</i> / <i>vi-</i> <i>vr-</i> > <i>vr-</i>	
<i>y</i> +							
<i>r</i> +							
Changes are presented in the form of Skt. cluster > G reflex Clusters in which significant changes occurred are highlighted in bold Note (1): Includes palatalisation of coronal stops.							

- Stop + nasal: $j\tilde{n}- > \tilde{n}-$

The OIA $j\tilde{n}$ cluster combines two homorganic occlusive consonants, and may have been pronounced as a unitary ‘pre-stopped’ nasal. This articulation was possibly retained in Gāndhārī, but without the orthographic j acting as a place marker for the nasal, and there was possibly little phonetic distinction between OIA $j\tilde{n}$ and Gāndhārī \tilde{n} (§ 8.3.1.2).

- Stop + y : $Ty- > T_p^{282}$

Word initial $Ty-$ clusters are not common in OIA (§ 8.3.1.4.3). This development is indicative of the weaker articulation of the y semi-vowel discussed above, which appears to have led to its lenition and deletion in word initial positions. These mostly involve coronal stops which are typically palatalised in the transition.

- Fricative + stop: $S\{T / T^h\}- > T^h-$

In most cases these clusters develop through a process of spreading of the place feature of the sibilant as a glottal gesture in aspiration of the stop (§ 8.3.3.1).

- Fricative + r : $\#sr- > \#ś-$

This development is suggestive of a weaker articulation of the r semi-vowel, similar to y , which is also observed in some of the medial clusters as discussed below.

8.3.6.2 Word medial clusters

The Gāndhārī reflexes of OIA medial two consonant clusters are summarised in Table 8.3.6.2. Clusters have been grouped in terms of combinations of manner, based on the MIA and Pali consonant strength hierarchies (§ 6.1.3), and changes are recorded in summary form to represent the most typical developments. In this table blank cells are those in which Sanskrit clusters are rare or do not occur, and changes are presented in the form of Skt. cluster $>$ G reflex. The reflexes listed represent the most likely developments interpreted from the data in the analysis sections. Where advanced reflexes occur alongside historic forms and the extent of the development in Gāndhārī is uncertain, both forms are listed and highlighted in round () parentheses. While progressive manner assimilation is more common in the later stage MIA cluster developments, this only occurs in some of the Gāndhārī reflexes, which are highlighted in bold in the table. Cases in which the record of Gāndhārī reflexes is very limited have mostly been excluded.

²⁸² The T_p- subscript indicates that the stop reflex was palatalised. Refer to the list of ‘Symbols’ section.

Table 8.3.6.2: Typical Gāndhārī reflexes of OIA word medial two consonant clusters

	Stop (T)	Nasal (N)	Fricative (S)	<i>l</i>	<i>v</i>	<i>y</i>	<i>r</i>
Stop (T) +	-T ₁ T ₂ - > -T ₂ T ₂ -	- <i>tm</i> - > - <i>tv</i> - - <i>jñ</i> - > - <i>ññ</i> -	-TS- > -TS- stable	-T <i>l</i> - > -T <i>l</i> - -T <i>v</i> - > -T <i>v</i> -		-T <i>y</i> - > -T <i>T</i> - ⁽¹⁾ -T <i>r</i> - > -T <i>r</i> -	
Nasal (N) +	-NT- > -NT- stable	- <i>nm</i> - > - <i>mm</i> - small sample	- <i>ṃS</i> - > - <i>ṃS</i> - small sample	-N <i>v</i> - > -N <i>v</i> - small sample		-N <i>y</i> - > -N <i>N</i> - ⁽¹⁾ - <i>mr</i> - > - <i>mr</i> - small sample	
Fricative (S) +	-ST- > (-ST- / -TT-) mixed -ST ^h - > (-ST ^h - / -TT ^h -) mixed	-SN- > (-SN- / -SS-?) uncertain - <i>Sm</i> - > - <i>Sp</i> - - <i>hm</i> - > - <i>mm</i> - / - <i>m^h</i> -	none in G	-S <i>l</i> - > -S <i>l</i> - rare <i>Sv</i> > <i>Sp</i>		-S <i>y</i> - > -S <i>S</i> - ⁽¹⁾ -S <i>r</i> - > (-S <i>r</i> - / -S <i>S</i> -?) mixed	
<i>l</i> +	- <i>l</i> T- > -TT-	small sample	none in G			- <i>ly</i> - > - <i>ll</i> -	
<i>v</i> +		none in G				- <i>vy</i> - > - <i>vv</i> - - <i>vr</i> - > (- <i>vr</i> - / - <i>vv</i> -?)	
<i>y</i> +							
<i>r</i> +	- <i>r</i> T- > (-TT- / - <i>r</i> T-) mixed	- <i>r</i> N- > -NN-	- <i>r</i> S- > (-SS- / - <i>r</i> S-) mixed	- <i>rl</i> - > - <i>ll</i> - - <i>rv</i> - > (- <i>rv</i> - / - <i>vv</i> -?)		- <i>ry</i> - > (- <i>ry</i> - / - <i>yy</i> -?) uncertain	
<p>Changes are presented in the form of Skt. cluster > G reflex</p> <p>Round parentheses () indicate alternate developments</p> <p>Changes involving progressive manner assimilation are highlighted in bold</p> <p>Note (1): Includes palatalisation of coronal stops.</p>							

Reflexes in each of the manner combinations are discussed in more detail below, categorised according to the types of assimilation which occurred. While the developments are considered primarily in terms of combinations of manner, concurrent place assimilation occurs in many cases within each manner grouping, and these patterns are discussed as appropriate.

8.3.6.2.1 No Change

Many OIA medial cluster combinations remained stable in Gāndhārī:

- Stop + fricative: $kṣ > kṣ$, $ts > ts$
- Nasal + Stop: $NT > NT$; Nasal + fricative $mś > m\{s, ś\}$
- Stop plus semi-vowel: $T\{l, v, r\} > T\{l, v, r\}$.
- Nasal + semi-vowel $Nv > Nv$, $mr > mr$,
- Semi-vowel + semi-vowel: $ry > ry / (yy)$, $vr > vr / (vv)$, $rv > rv / (vv)$

Apart from the final case these all occur in clusters of equal or reducing strength, based on the assumed Pali strength hierarchy (§ 6.1.3).

8.3.6.2.2 Place assimilation only

Regressive place assimilation is seen between medial consonants of similar manner:

- Between occlusives: Stop + stop: $T_1T_2 > T_2T_2$, Nasal + nasal: $nm > mm$
- Between approximants: $rl > ll$. Also progressive: $ly > ll$, $vy > vv$

Place assimilation between pairs of stops and nasals follows expected patterns of anticipatory assimilation of place by the first consonant in favour of the second. This occurs in all cases regardless of the combinations of place, so that first placed labials and velars, which are expected to be more robust (§ 6.1), assimilate to following coronals, as in $-kt- > -tt-$ and $-pt- > -tt-$ for example.

Since the OIA semi-vowels are all continuants it might be expected they would follow similar patterns of regressive place assimilation. Both regressive and progressive assimilation occur, however, as for example in $rl > ll$ and $ly > ll$ respectively, and the differences between the semi-vowels may be more correctly treated as subtle differences in manner due to the variations in the shaping of the tongue during articulation (§ 6.1.3). These clusters are also considered below in terms of the relative strength hierarchies of $l > v > y > r$, where the cluster assimilates in favour of the stronger consonant.

8.3.6.2.3 Manner assimilation

Manner assimilation is both regressive and progressive, and in many cases involves concurrent assimilation of place in favour of one of the consonants.

Regressive manner assimilation

Place preserved in homorganic clusters – manner assimilation only:

- Stop + nasal: $j\tilde{n} > \tilde{n}\tilde{n}$,
- Semi-vowel + nasal: $r\eta > \eta\eta$

Place assimilated to that of the dominant consonant – manner and place assimilation:

- Semi-vowel + stop: $l\{p, bh\} > \{pp, bbh\}$. Dental approximant l to labial stop p / b .
- Semi-vowel + nasal: $rm > \eta m$. Retroflex / alveolar continuant r to labial nasal m .
- Fricative + nasal: $hm > mm / m^h$. Glottal fricative h to labial nasal m .
- Semi-vowel + semi-vowel: $rl > ll$. Alveolar tap to dental lateral approximant.

Apart from the $j\tilde{n}$ cluster between two occlusives, and the rl cluster between two approximants, these all involve assimilation of both manner and place in favour of the second consonant, which also has the greater degree of occlusion or strength. The $rl > ll$ result suggests that l is slightly more occlusive than r .

Mixed results are also seen in a number of clusters which are sometimes preserved, but also develop through similar patterns of regressive assimilation of both manner and place in favour of the second consonant. The following cases have been identified in round parentheses () in Table 8.3.6.2:

- Fricative + stop: results are varied
 - Coronal homorganic unaspirated clusters tend to be preserved, $ST > ST$, except for:
 - Homorganic palatals tend to assimilate: $\acute{s}c > cc$
 - Velar stops assimilate sibilants: $\acute{s}k > kk$
 - Place aggregation of sibilants: $Sp > \acute{s}p$. (Small sample)
 - Coronal homorganic aspirated stops are both preserved and assimilated, although some results are uncertain: $\acute{s}th > \acute{s}th$, $sth > th$
- Semi-vowel + stop:
 - Retained orthographically in many cases: $rT > rT$
 - Otherwise assimilated across all places: $rT_1 > T_1T_1$

- Semi-vowel + fricative:
 - Retained orthographically in many cases: $rS > rS$
 - Homorganic $r\varsigma > \varsigma\varsigma$ assimilated. Palatal sibilants show a tendency to place assimilation by retroflex r : $r\acute{s} > \{\acute{s}\acute{s}, \varsigma\varsigma\}$.
- Semi-vowel + semi-vowel:
 - Some results are uncertain. Clusters are both preserved and assimilated: $ry > ry / (yy)$, $rv > rv / (vv)$

Assimilation mostly follows a pattern of regressive assimilation of manner, along with concurrent assimilation of place in non-homorganic clusters, as in $rk (> \acute{t}k) > kk$, or $rt (> \acute{t}t) > tt$ for example, where r may become \acute{t} prior to place assimilation by the stop. In the second case $\acute{t}t$ reflexes are also recorded, supporting the concept of an intermediate stage, developing in this case through progressive assimilation between the coronal stops²⁸³.

Regressive manner assimilation of a more open initial consonant has been interpreted in terms of strength-based manner assimilation in Pali (§ 9.3), which can also be explained in terms of anticipatory assimilation of the greater degree of occlusion of the second consonant by the first. The assimilation tendencies in the $ry > ry / (yy)$ and $rv > rv / (vv)$ clusters suggests that both y and v are both slightly more occlusive than r .

The mixture of conservative and innovative forms in the reflexes of some clusters, such as $rT > (TT / rT)$ and $rS > (SS / rS)$ for example, compares with the full assimilation of these clusters in later stage MIA languages, suggesting that the conservative forms in Gāndhārī are less likely to be borrowings. One explanation for the mixed results is that archaic speech patterns and historical spellings remained embedded in the language, at least to some extent, while advanced forms reflect synchronic phonetic developments that were occurring in the language. The continued use of both reflexes over several centuries could also reflect spelling alternations which developed during the early stages of the sound transition, or even reflect dialectic divergences, while both forms were subsequently retained and used freely in the diachronic record. In either case the mixed developments point towards a more gradual progression in these changes in the Gāndhārī language and orthography than appears to have occurred in the other MIA languages.

²⁸³ This is similar to the process described by Gordon (2016: 132).

Progressive manner assimilation

This occurs in the following cases with the semi-vowel *y*, which are marked in bold in Table 8.3.6.1:

- Stop, Nasal or fricative + *y*: {T, N, S}*y* > TT, NN, SS respectively. These often involve concurrent palatalisation of coronal consonants in the geminate, as in *ty* > *cc*, *ny* > *ññ* and *śy* > *śś* for example.

These all involve progressive assimilation of the semi-vowel by the more occlusive stop, nasal and fricative consonants. These reflexes are typically recorded as single unweakened consonants in the Kharoṣṭhī script, which are interpreted as geminate reflexes for the reasons explained in each case (§§ 8.3.1.1, 8.3.1.2, 8.3.3.4.3), as well as through comparison with other MIA languages (§ 9.1.2). Progressive manner assimilation of more open final consonants has been interpreted in terms of strength-based manner assimilation in MIA and Pali (§ 9.3.1), and has also been explained in terms of the diachronic development of Sanskrit gemination patterns as a strengthening or adjustment of sonorities across the syllable boundary (§ 9.3.4).

- Semi-vowel clusters: *ly* > *ll*, *vy* > *vv*.

These reflexes follow the general pattern of assimilation of following *y* in Gāndhārī, and provide support to the suggestion that both *l* and *v* are slightly more occlusive than *y*, as discussed in § 6.1.3.

Mixed results are also seen in a small number of other clusters which are sometimes preserved, but also develop through similar patterns of progressive assimilation of both manner and place in favour of the second consonant.

The following cases involving the semi-vowel *r* are also marked bold in the table:

- Fricative + *r*:
 - Retained orthographically in many cases: {ś, s}*r* > {ś, s}*r*
 - Otherwise, assimilation by the sibilant in manner and place {ś, s}*r* > {śś / ss}
- Semi-vowel *v* + *r*:
 - *vr* > *vr* / *vv* possibly assimilates in favour of *v*, confirming its relative strength compared with *r*.

Reflexes in the following case are less certain:

- Fricative + nasal:
 - Orthographic \bar{s} and \bar{z} are interpreted to represent retained homorganic clusters, as in $\bar{s}\eta > \bar{s}\eta$ and $\bar{z}\eta > \bar{z}\eta$, although these may also represent geminate sibilants.

8.3.6.2.4 Lenition / fortition

The following cases involve either lenition of the second placed labial consonants, as in $m > v$, or fortition, as in $m > p$ and $v > p$. At the same time, there is no assimilation of the cluster in favour of either consonant.

- $tm > tv$ (a small number of tt and pp reflexes are also recorded, which are suggested as borrowings)
- $Sm > Sp$, $Sv > Sp$

8.3.6.3 Three consonant clusters

OIA three consonant clusters typically occur in medial positions, and their development in Gāndhārī appears to vary according to the syllable division in the OIA clusters (§ 8.3.5). Once the OIA syllable divisions are considered, the two consonant components of three consonant clusters tend to follow the same development pattern as the same stand-alone two consonant clusters. The N.T{S, R} and N.ST clusters show the most consistent developments, while the S.TR clusters are also reasonably consistent in their changes, with both retaining the original OIA syllable structure.

8.3.6.4 Scribal variations

It has been difficult to discern any consistent patterns of difference in the orthographic habits of individual scribes, or within or between groups of scribes. Such patterns are more likely to be evident in reflexes of the ST, SN, SR and r C type clusters, where the overall record is most variable. Some of the developments in which variable reflexes are recorded for these clusters are reviewed below in relation to the patterns of scribal use:

Reflexes of ST clusters:

- $-ṣṭh-$ > $-ṭh-$ / $-ṭh-$ mixed usage by BL scribe 1 and BC scribe 5.
- $sth-$ > $th-$ / $ṭh-$ / $ṭh-$ / $ṭ-$ mixed usage by BL scribes 1 and 4 and the RS scribe.

- *-sth-* > *-ṭh-* / *-ṭ-* mixed usage by BL scribes 1 and 4 and the RS scribe. Although there is some intermixing of their use, the former reflex appears to have been used more frequently by the BL scribes, and the latter more by the RS scribe, suggesting this could be a matter of local scribal conventions (§ 8.3.3.1).

The recording of mixed reflexes among the ST clusters appears to be the result of confusion about the changes and how to represent them, indicating some form of allophonic or transitional pronunciation.

Reflexes of SN clusters:

- *-ṣṇ-* > *-ṣ-* / *-ṣ̄-* mixed usage by BL scribes 2 and 4, while *rṇ* is used by BL scribe 1.
- *-śm-* > *-ṣp-* / *-śp-* / *-sp-* mixed usage by the RS scribe.
- *sm-* > *sp-* / *śp-* / *sv-* by different BL scribes.
- *-sm-* > *-sp-* / *-(m)m-* by BL scribe 1 in reflexes of the Sanskrit *-asmin* pronominal locative case ending, as well as the verbal form *asmi*. *sp* appears to have been preferred by the RS scribe and BL scribe 9, while *s(s)* was also used by both scribes.

Reflexes of SR clusters:

- *sv-* > *sp-* / *śp-* mixed usage seen in both the BL and RS manuscripts.
- *-sy-* > *-s(s)-* / *-ṣ-* mixed usage by the RS scribe in the genitive singular endings *-asya* > *-asa* / *-aṣa*.
- *śr-* > *ś-* / *ṣ-* mixed usage with the former used mostly by RS scribe, and the latter, seemingly assimilated by the prior *r* sound, by the BL scribes.
- *-śr-* > *-śr-* / *-ś(ś)-* / *-ṣ(ṣ)-* mixed usage by the RS scribe and among the BL scribes.
- *-sr-* > *-sr-* / *-s(s)-* / *-ṣ-* mixed usage by the RS scribe, with *s(s)* and *ṣ* used by BL scribes.

Reflexes of rC clusters. Mixed reflexes are recorded by most scribes in the following clusters unless otherwise indicated:

- *-rt-* > *-(t)t-* / *-(ṭ)ṭ-* / *-(ṭ)ṭh-*.
- *-rth-* > *-rth-* / *-(t)th-* in reflexes of the common Sanskrit word *artha-* > G *artha-* / *atha-*.
- *-rm-* > *-rm-* / *-(m)m-* in reflexes of the common Sanskrit words *karman-* > G *karma-* / *kama-*, and *dharmā-* > *dharmā-* / *dhama-*.

- $-rś-$ > $-rś-$ / $-(ś)ś-$ / $-(ś)ś-$.
- $-rṣ-$ > $-rṣ-$ / $-(ṣ)ṣ-$ by the BL scribes 1 and 2.
- $-rv-$ > $-rv-$ / $-(v)v-$ in reflexes of the common Sanskrit word *sarva-* > G *sarva-* / *sava-*.
- $-ry-$ > $-ry-$ / $-ri-$.
- Metathesis of rC clusters is also common in the Aśokan inscriptions (§ 8.3.4.4.3).

While on the one hand these variable reflexes suggest a divergence of orthographic habits within and between the scribal communities, there also appears to have been much shared knowledge and use of the variable forms. As discussed in § 8.2.5.5, this points towards significant levels of interaction between the scribal communities, in order, for example, to establish a best graphic representation of phonemes and clusters in transitional stages of language change. This suggests that significant interaction must have occurred between the various Gāndhāran scribal groups, whether through synchronic or diachronic relationships within and between communities, in contrast to isolated or siloed communities developing their own distinct practices. At the same time, alternations in the use of many of the graphic variants, both by individual scribes and within communities, suggest that there was a degree of uncertainty or lack of agreement about the correct forms of representation.

The following examples of some commonly adopted forms are also indicative of shared orthographic habits among the communities :

- $-kṣ-$ > $-kṣ-$ / $-(k)kh-$, where the latter form is only found in reflexes of technical terms, such as Skt. *bhikṣu-* > G *bhikhu-*, which appear to be borrowings from an eastern dialect, and copied by both the BL and RS scribes. Scribes have typically retained the unchanged form in all other words.
- $pr-$ > $pr-$ / $p-$, where the latter reduced form is typically recorded by all scribes in reflexes of the Sanskrit verbal prefix *prati-* > G *paḍi-*, although the BL scribe 1 has also used the retained form *prati-* in many cases.
- $-ṁT-$ > $-T-$ / $-(ṁT-)$. While the marking of *anusvāra* appears have been a common practice in the third century BCE Aśokan inscriptions, it is mostly unrecorded by the BL and RS scribes. While this could indicate the loss of preconsonantal nasality as much as a graphic practice, the habit of not marking *anusvāra* had become widespread and shared among the first and second century Buddhist scribal communities.

8.3.7 A record of language change in progress

Similar to the variable orthographic record of reflexes of the OIA single intervocalic consonants, which includes both historic and advanced orthographic forms, along with diacritic marked *akṣaras* representing transitional forms (§ 8.2.1.2.6), the Gāndhārī reflexes of OIA consonant clusters include a similar range of variable forms. These graphemic variations most likely reflect actual variation in the pronunciation of the language, offering further insights into the patterns of sound and language change in progress during the period covered by the source materials. Examples include:

- The mixed orthographic record for many cluster reflexes, including the $ST > ST / TT$, $ST^h > ST^h / TT^h$, $Sr > Sr / SS$ and $vr > vr / vv$ clusters, along with preconsonantal *r* clusters including $rT > TT / rT$, $rS > SS / rS$, $rv > rv / vv$, $ry > ry / yy$ (Table 8.3.6.2). In many instances, such as for the *ST* clusters (§ 8.3.3.1), graphemes representing both conservative and innovative forms are used inconsistently by the scribes, making it difficult to discern the nature of the changes that had taken place. In other instances, such as for many of the *rC* type clusters (§ 8.3.4.4) it is unclear whether assimilated reflexes represent borrowings from source dialects or the actual development, while retained forms may represent habitual historic spellings or dialectical divergences.
- The use of diacritic marked reflexes whose intention remains unclear, suggesting that the scribes were dealing with how best to record variable sounds in the language. These include examples such as $sth > th$, which may represent transitional stages in the assimilation of this cluster (§ 8.3.3.1), and $ṣṇ > ṣ$, and $\{śn, sn\} > ṣ$, which, in the absence of a conjunct grapheme in the Kharoṣṭhī script possibly represents the retained cluster (§ 8.3.3.2).

Furthermore, the general assimilation of *y* in *Cy* clusters, while *r* is only partially assimilated in *Cr* clusters and *Cl* and *Cv* clusters are generally retained in Gāndhārī, indicates a partial stage of development of the full assimilation of semi-vowels by preceding consonants. This compares with the full assimilation of these clusters in the later stage MIA languages and Pali²⁸⁴ (9.2.2), indicating a partial stage of development in Gāndhārī where assimilation of the semi-vowels in clusters was in transition. This sheds light on the progressive stages of these changes as well as the relative strengths of the semi-vowels in clusters, as discussed further in § 9.3.

²⁸⁴ These developments may be more structured in Pali, as discussed in § 9.3.5.

8.4 Syllable patterns in Gāndhārī

As discussed in § 4.5, the syllable patterns in Sanskrit are determined by poetic prosody as well as the relative sonority profiles of consonant clusters. In this context, it is difficult to assess the changes to these patterns in a vernacular MIA language like Gāndhārī, which emerged as the dominant dialect of the Gandhāra region before being adopted as the ecclesiastical language of Buddhism in the area (§ 2.1). While the nature of the relationship between the syllable structures of poetic Sanskrit and those of the spoken OIA dialects is uncertain, this is even more so for the spoken forms of Gāndhārī, while any metrical patterns in MIA and Pali source texts do not appear to have been fully incorporated into the Gāndhārī texts (§ 8.1). While making direct comparisons with changes from the syllable structure of poetic Sanskrit is difficult for these reasons, a preliminary assessment of some of the apparent and emerging syllable patterns in Gāndhārī is included in this section. These include the effects of gemination of consonant clusters and the weakening of intervocalic consonants, while changes to vowel patterns are also addressed. A short section on the syllable patterns in Torwali is also included.

8.4.1 Changes to consonant clusters – geminate reflexes

The effects of gemination vary according to the type of consonant cluster, and the recording of variant historical and progressive forms of reflexes assists in revealing the patterns of change in progress.

Where historic clusters have decreasing sonority, these are typically divided as coda and onset, and gemination does not affect this pattern since geminates also divide in the same way. This is seen in the following examples, which compare words for which both historic and advanced forms of the original cluster are recorded. In these cases, gemination develops through regressive manner assimilation of the first consonant: G *ar.tha* and *at.tha*- (Skt. *artha*-); G *ca.tur.di.śa*- and *ca.tud.di.śa*- (Skt. *cā.turdi.śa*-); G *dar.śana*- and *daś.śa.na*- (Skt. *darśana*-); G *dhar.ma*- and *dham.ma*- (Skt. *dharma*-).

Where historic clusters have increasing sonority, these are typically considered as cluster onsets to the following syllable, and gemination results in a distribution as a coda and onset. Assuming this was not already considered for metrical reasons, this redistribution adds weight to the prior syllable. In these cases, gemination develops

through progressive manner assimilation of the second consonant: This is seen in the following examples: G *ca.kṣu-* and *ca.k.khu-* (Skt. *caḥṣus-*); G *a.di-kra.mo* and *a.bhik.ka.tu* (Skt. *ati-* / *abhi-krama-*); G *sa.mya-* and *sam.ma-* (Skt. *saṃya-*); G *sa.ha.sra-* and *sa.has.sa-* (Skt. *sahasra-*); G *-a.sya* and *-as.sa* (Skt. *-aśya*); *pa.ri.vra.ye* and *pa.riv.va.ye* (Skt. *parivrajat*).

8.4.2 Changes to single consonants

The lenition of intervocalic consonants does not affect the syllable structure, as in Skt. *acala-* > G *ayalu-*, *katara-* > *kaḍara-* and *apara-* > *avara-* for example. In cases where the weakening leads to loss of the consonant, the number of vowels appears to be retained in most cases, supporting a similar syllable structure, as in Skt. *anarthikaḥ* > G *aṇathio*, *kapālakena* > *kavalaena*, *-nikāyam* > *-ṇiao* and *anupādāna-* > *aṇuadana-* for example. As discussed in § 8.2.1.2.1 and § 8.2.4.2, recording of the vowels without the consonants suggests a weakened and barely audible pronunciation as *ya śruti*. This was realised phonetically in the vowel hiatus, and it appears that the scribes did not consider it necessary to record this.

8.4.3 Changes to vowels and vowel sequences

As evidenced by the literary Prakrits, most of the weakly articulated single consonants were eventually lost in the MIA languages, leading to vowel sequences that eventually contracted to single long vowels and diphthongs, with subsequent reduction in syllable numbers (Misra 1967: 166, 204). Although not recorded consistently, changes such as reduction of the sequences *aya* > *ai* > (*ai*) > *e* and *ava* > *ai* > (*au*) > *o* (e.g., Allon 2001: 73-74, 77-78) are widely attested in the first and second century CE Gāndhārī manuscripts, indicating that processes of vowel deletion were beginning to occur in Gāndhārī, at least in situations where the vowel merger is phonetically straightforward. On the other hand, and although occurring less frequently, more complex vowel sequences involving three vowels, such as BHS *plṭhakaḥ* > G *priao* (Glass 2007: 246), and Skt. **udariyakam* > G *udario* (Salomon 2008: 426), were also retained in Gāndhārī.

Another area where syllable structures are altered by changes to consonant clusters follows the resolution of clusters using epenthetic vowels. Depending upon how the phonological record is interpreted, the recording of epenthetic vowels represents the

creation of additional syllables. This is seen in the following examples: Skt. *glā.naḥ* > G *[ghi.l]/(*a).no*; Skt. *u.pa.kle.śān* > G *u.va.ki.le.śa*; Skt. *dve* > G *du.ve*; Skt. *kva.cit* > G *ku.va.ci*; Skt. *ślo.ke.na* > G *śi.lo.ga.no*; Skt. *Bā.hṛ.ka-* > G *Ba.ha.li.a*; Skt. *pad.ma* > G *pa.du.ma*; and Skt. **tan.va.ka-* > G *ta.nu.va.ka-*. A collation of all cases of epenthesis recorded in the Gāndhārī source materials is presented in § 10.1, including review of their phonetic and phonological features along with common conditioning factors.

8.4.4 Syllable patterns in Torwali

It has been observed that tautosyllabic consonant clusters are not permitted in Torwali, and consonant clusters are only allowed across syllable boundaries (Lunsford 2001: 20, 33-35). Hence clusters of the type *-r.T-*, *-r.S-*, or *-S.T-*, in which sonority falls across the syllable boundary, can occur, while clusters involving the most sonorant glides *y* and *v* as second members do not occur since these can only form tautosyllabic onset clusters with any prior consonant having less sonority. The partial retention of the heterosyllabic *rT* and *rS* clusters and total assimilation of the *Cy* clusters in Gāndhārī suggest an affinity with these patterns, where assimilation of the *-r{ T, S }-* clusters may have developed more slowly due to a larger areal association with, or even dialectic borrowings from, the Dardic languages. On the other hand, the tautosyllabic *Cr* type clusters are also commonly retained in Gāndhārī, suggesting an overall slower pace of change in the Gāndhārī clusters, possibly due to its relative isolation from the other MIA languages (§ 9.3.1), as much as any influences from the Dardic languages (§ 8.2.6.3).

8.5 Retroflex consonants in Gāndhārī

The retroflex consonants arose in OIA under a variety of assimilation conditions, along with the influence of later borrowings and analogical spread²⁸⁵. In the development of MIA languages such as Gāndhārī, in which extensive sound changes occurred, many of the elements which conditioned the spread of retroflexion may have changed in their own right, leading to the potential loss of these phonetic influences on other consonants. These changes occurred to both single consonants as well as in some of the consonant clusters.

As discussed in § 8.2.2.2.2, there is a widespread intermixing in the use of both the dental and retroflex nasals *n* and *ṇ* in the Gāndhārī source materials, regardless of their original distribution, to the extent that one or the other symbol is used almost exclusively by particular scribes. Hence although the phonetic factors for the conditioning of *ṇ* are still present in many Gāndhārī words, it appears to have been recognised that the distribution of the nasals as *n* or *ṇ* did not distinguish minimal pairs and had little if any practical lexical implications in the language. On the other hand, the distribution of the retroflex sibilant *ṣ* and the stops *ṭ* and *ṭh* in Gāndhārī reflexes appears to be more consistent with their original OIA representation, and can be examined with more certainty on that basis. The retroflex tap *r* is a common conditioning element of both *n* and *s*, acting both remotely as a preceding single intervocalic consonant, and also as the first element in clusters. While intervocalic *r* is typically retained (§ 8.2.4.2), when it occurs as the initial element in clusters it is often assimilated (§§ 8.3.4.4.2, 8.3.4.4.3), as seen in the following examples of *-rṣ-* clusters:

- Skt. **pārṣamda* > G *paṣamda*, *vyākārṣīt* > *viaghaṣe*, *varṣāvāsaḥ* > *vaṣavaso*, *caturṣu-* > *caduṣu*, where the *ṣ* reflex is assumed to represent the geminate (*ṣṣ*) (Table 8.3.4.4.3a).

There are also many cases in which the conditioning elements are retained and there is no change in the recording of *ṣ*, as seen in the following examples:

- Skt. *skandheṣu* > G *kadheṣu*, **dveṣaḥ* > *doṣa*, *puruṣaḥ* > *puruṣa*, *-viṣāṇakalpo* > *-viṣaṇagapo*, *-viṣūkāni* > *-viṣoāni*, *asrauṣīt* > *[a]śroṣi*, *tuṣitāḥ* > *tuṣida*, *nyaṣīdat* > *ṇiṣīdi*, *viṣeṣa-* > *viṣeṣa-*, *maharṣeḥ* > *mahirṣ[i]ṇa* (§ 8.2.3.2). In many of these cases

²⁸⁵ Refer to Appendix A1 for a more detailed account of the OIA retroflex consonants.

the mid-high vowel *e* appears to act as a conditioning influence in both Sanskrit and Gāndhārī.

Retroflex *ṣ*, whether conditioned or non-conditioned itself, also influences any following dental stop *t(h)* in Sanskrit clusters, but is typically assimilated in the Gāndhārī reflexes (§ 8.3.3.1):

- Skt. *aṣṭa-* > G *aṭha-*, *śreṣṭha-* > *sreṭha-*, *nirdiṣṭaḥ* > *ṇidiṭhu*, *prṣṭaḥ* > *proṭhu*, *akṣṣṭaḥ* > *akroṭha-*, *drṣṭa-* > *driṭha-*, *tiṣṭhiti* > *tiṭhiti*, *kaniṣṭha-* > *kaniṭha-*, *jyeṣṭha-* > *jeṭha-*, *suṣṭhu-* > *suṭhu-*, where the *ṭh* reflex is assumed to represent the geminate (*t*)*ṭh* (Table 8.3.3.1a). In many of these words, other retained phonetic elements such as the high vowel *i* may also contribute to a continued retroflex pronunciation of *-ṭh-*.

The recording of Skt. *bhāṣita* ($\sqrt{bhāṣ}$ “say, speak” ppp. [MW s.v.]) > G *bhaṣide* (Marino 2017: 126) is an example of non-conditioned OIA *ṣ* also being retained in Gāndhārī.

These examples demonstrate that although the original phonetic conditions under which they arose are no longer present in many instances, the retroflex consonants are still recorded as such. This raises wider questions about the extent to which the possible loss or retention of assimilated features may have occurred, and how the language may have adapted to these changes. By way of comparison, retroflex *ṣ*, *ṭ* and *ṭh* are recorded as phonemes in Torwali (Lunsford 2001: 22), suggesting the language may have preferred to retain these sounds for any additional contrastive patterns they offered, even as the conditioning influences were lost in some cases. On the other hand, internal changes to the conditioning of *ṣ* may have played a part in the phonemic merger of the sibilants in other MIA languages (§ 9.2.3), even though these were retained in Gāndhārī, possibly due to areal influences from the northwestern languages (§ 8.2.6.3). In either case it is not possible to address these wider questions with any certainty, other than to suggest that the changes to the conditioning influences documented in the Kharoṣṭhī record offer further insight into these patterns of sound change in progress.

8.6 Consonant sandhi phenomena in Gāndhārī

The Sanskrit sandhi system formalises the pronunciation and recording of some of the allophonic variations of consonants and vowels according to the phonetic environments in which they occur (§ 4.2). While sandhi is generally not reflected in such a way in written Gāndhārī (Salomon 2000: 78), some vowel sandhi phenomena have been observed in most of the manuscript source materials²⁸⁶. Examples of consonant sandhi also appears to be sparse, and many of the Sanskrit applications appear to have limited relevance in Gāndhārī where the overall patterns of consonant use have changed significantly. For example, word final stops are relatively infrequent in OIA declensional and conjugational endings, and when they do occur are typically lost in Gāndhārī, so that consonant to consonant junctions between words and members of compounds do not arise (e.g., §§ 8.2.1.3, 8.2.1.1). Some important applications of internal sandhi also appear to have limited relevance in Gāndhārī. One instance is where consonant junctions occur between preverbs and stems, as in Skt. *ut + gama-* > *udgama-* (§ 4.2) for example. Such a medial consonant cluster would typically assimilate to a geminate in Gāndhārī, as in Skt. *udgama-* > G *u(g)gama-* and the original consonant sandhi is merged within the ongoing language change. Other examples relate to changes in the patterns of retroflex consonants discussed in the previous section, such as variable use of the characters for the dental or retroflex nasals by different Gāndhārī scribes, which represents a breakdown of the formal recording of *ṇ* as a spoken allophone of *n* in certain environments under the Sanskrit sandhi rules. A similar example occurs in the case of the retroflex sibilant *ṣ*, which is retained in some cases even though the conditioning element, such as *r* in *rṣ* clusters, is assimilated, as in Skt. *varṣāvāsaḥ* > G *va(ṣ)ṣavaso*. Similarly, in cases such as *nirdiṣṭa-* > *ṇidi(t)thu-*, the retroflex *ṣ*, which is conditioned by the preceding *i* and subsequently conditions the following stop in Sanskrit, becomes assimilated in Gāndhārī. Although not prescribed in the Sanskrit sandhi system, the retroflex stop is seemingly retained in response to the preceding high vowel. Such an approach is consistent with the recording of consonant allophones for the intervocalic stops for example (§ 8.2), suggesting that many apparent cases of sandhi most likely represent the variable recording of allophones rather than the application of a formal system similar to Sanskrit.

²⁸⁶ E.g., Salomon (2000: 78), Allon (2001: 99-101), Glass (2007: 124), Salomon (2008: 127-128), Baums (2009: 188- xx), Silverlock (2015: 274), Schlosser (2016: 90), Marino (2017: 133).

8.7 Later stage developments: third century CE inscriptions and manuscripts

The available Gāndhārī source materials dated to the third century CE are of limited extent, and include a small number of Kharoṣṭhī inscriptions belonging to the later Kanīṣka era, along with numerous fragments of palm leaf manuscripts found in Bamiyan, held in the Schøyen, Hirayama and Hayashidera collections (§ 3.3). While these manuscripts are in a very fragmentary condition, some important Buddhist texts have been identified among those held in the Schøyen collection, and although partial in extent, details of these texts have been published in several sources (§ 3.5).

While both the phonology and morphology of the language used in these manuscripts mostly reflects the Gāndhārī represented in the first and second century CE BL and RS manuscripts, a number of conservative Sanskritic features have been restored to the written form of the language, in a process described as ‘Sanskritisation’²⁸⁷. Compared with the more cursive script seen in some of the BL and RS birch bark manuscripts, the Kharoṣṭhī characters used on these manuscripts appear to be more carefully written and better aligned, which is possibly due in part to the limited vertical dimension and long horizontal format of palm leaf segments. While not applied consistently throughout the manuscript fragments, changes to the consonant phonology include the following adjustments to conjunct character reflexes, which are listed by way of example:

- Gāndhārī spellings in the MPS-G^B manuscript include *vyakta*, *sapta*, *puṣkirini* and *brahme*, where the clusters are normally reduced to geminates (§§8.3.1.1, 8.3.3.1, 8.3.3.2), *ayusmato* where the cluster is normally rendered as *-sp-* (§ 8.3.3.2), and *sthasyamti* and *sthuvaṃ* where the word initial sibilant is normally deleted ahead of the aspirated stop (§ 8.3.3.1)²⁸⁸.
- Gāndhārī spellings the Bhk-G^B manuscript include the use of similar uncommon conjunct characters such as *ṣka*, *hma* and *ṣya* where the clusters are normally reduced to geminates (§ 8.3.3.1, § 8.3.3.2, § 8.3.3.4.3), *tma*, which is normally reduced to *tv* (§ 8.3.1.2), and *smi* where the cluster is normally rendered as *spi* (§ 8.3.3.2) (Baums, Glass and Matsuda 2016: 186-187).

²⁸⁷ Allon and Salomon 2000: 268; Salomon 2001.

²⁸⁸ Allon and Salomon 2000: 266-269; Salomon 2001: 243-247.

- Although the overall degree of Sanskritisation in the Bspt-G^B manuscript appears to be less than in other Bamiyan Gāndhārī documents, some restored conjunct characters such as *sya* and *tma* are recorded in the manuscript (Baums, Braarvig, Lenz, Liland, Matsuda and Salomon 2016: 268-269).
- Similarly, the overall degree of Sanskritisation in the EĀ-G^B manuscript appears to be less than in other Bamiyan Gāndhārī documents, although some restored conjunct characters such as *sya*, *vy* and *ly* are recorded in the manuscript (Jantrasrisalai, Lenz, Qian and Salomon 2016: 12-13).

These variants also provide examples of otherwise infrequent consonantal ligatures used for the above conjunct characters in the orthography, which have been included in the tabulation in § 5.3.4. Similar patterns of conservative phonology are also seen in some of the late second and early third century inscriptions of the Kaniṣka era, such as the use of *ky*, *thy*, and *py* on the Wardak vase and *ty* on the Tor Dherai potsherds (§§ 5.3.2.1, 8.3.1.4.3), and the retention of *rT* clusters such as *rg*, *rgh*, and *rd*, where geminate reflexes are otherwise expected (§ 8.3.4.4.1).

These adjustments in the written language shows that the Gāndhārī scribes were to some extent aware of and in touch with the contemporary Sanskrit traditions of the area by the third century CE, although the motivations for adopting these changes in Gāndhārī texts are uncertain. It is likely, however, that these changes reflect the growing influence of Sanskritic intellectual culture on written Buddhist documents which occurred around this time²⁸⁹. It is also unlikely that the conservative orthography adopted in these scholarly documents reflected the true sounds of the vernacular language, and that a diglossic situation was present in the language at this time (Jantrasrisalai, Lenz, Qian and Salomon 2016: 13). In the sense that these changes appear to have been a scholarly adaptation of the written language, they can be considered to some extent artificial and not a true reflection of the spoken language.

²⁸⁹ Salomon 2001: 248-251; Strauch 20012b: 149-152.

9. Comparative developments between Gāndhārī, MIA and Pali

9.1 Word initial consonant clusters

Word initial single consonants are generally stable in Gāndhārī (§§ 8.2.1.1, 8.2.2.1, 8.2.3.1, 8.2.4.1) and Pali (Oberlies 2019: 114-115), and significant changes only occurred to consonant clusters in initial positions. Changes that occurred in the word initial clusters recorded in the first and second century Gāndhārī Buddhist manuscripts, Pali, and the later period MIA literary Prakrits as represented by Śaurasenī ²⁹⁰, are listed comparatively in Table 9.1.1. Changes are presented in the format of Skt. cluster > MIA reflex. The developments in Pali and Śaurasenī are essentially the same and are listed together.

Table 9.1.1: Comparative development of Gāndhārī (G), Pali (P) and Śaurasenī (Ś) initial consonant clusters (The P and Ś cases are listed in the shaded rows)

		Stop (T)	Nasal (N)	Fricative (S)	<i>l</i> <i>v</i>	<i>y</i> ⁽¹⁾ <i>r</i> ⁽²⁾
T +	G:		<i>jñ-</i> > <i>ñ-</i>	<i>kṣ-</i> > <i>kṣ-</i>	Tl- > Tl- / Tr- Tv- > Tv-	Ty- > T- Tr- > Tr-
	P & Ś:		<i>jñ-</i> > <i>ñ-</i>	TS- > T ^h - ⁽³⁾	T{l, v}- > T-	T{y, r}- > T-
N +	G:					
	P & Ś:				<i>ml-</i> > <i>m-</i>	<i>mr-</i> > <i>m-</i>
S +	G:	ST- > T^h-	SN- > SN-		Sv- > Sv- / Sp-	Sy- > Sy- Sr- > S-
	P & Ś:	ST- > T ^h -	SN- > N ^h -		S{l, v}- > S-	S{y, r}- > S-
v +	G:					vy- > vy- / v- vr- > vr-
	P & Ś:					vr- > b-

Changes are presented in the form of Skt. cluster > G or MIA reflex

Clusters in which significant changes occurred in Gāndhārī are highlighted in bold

Notes: (1) palatalisation is often associated with *y*.

(2) cerebralisation of dentals is often associated with *r*.

(3) *kṣ-* > *kh-* East, *ch-* West

Source: compiled from Table 8.3.6.1, Misra (1967: 124, 130-136, 139) and Oberlies (2019: 142-146, 149).

²⁹⁰ The literary Prakrits, dating from around the middle of the first millennium CE, provide an indication of the progressed MIA developments. Śaurasenī developed from the central MIA Aśokan Prakrit, while also having many similarities with the eastern Prakrit, and is considered one of the earliest of the literary Prakrits (§ 2.2.3). Its development has been documented in a clear and accessible format by Misra (1967).

The main points of similarity and difference between these developments are as follows:

- In the Pali and Śaurasenī reflexes the OIA word initial clusters are reduced to single consonants in most cases. This suggests a process of lenition and deletion of the cluster medials which are typically weaker in the word initial OIA clusters. Exceptions to this pattern are seen in the following clusters, involving alternative processes:
 - $jñ- > ñ-$ which the initial stop is lost
 - $ST- > T^h-$ involving spreading of the fricative features from the sibilant to the stop
- The clusters are retained in most cases in the Gāndhārī reflexes. The exceptions to this pattern include the $jñ- > ñ-$, $Ty- > T-$, $ST- > T^h-$ and $Sr- > S-$ cluster developments, which are highlighted in bold in the table.

The changes to some but not all of the Gāndhārī word initial clusters suggest an intermediate stage of MIA development.

9.2 Word medial consonants and consonant clusters

9.2.1 Single consonants

The general patterns of change that occurred to the OIA intervocalic single consonants in the MIA languages are summarised below. A more detailed comparative tabulation of the reflexes is included in Table C1 in Appendix C.

- In the third century BCE Aśokan Prakrits some lenition of the intervocalic stops is recorded, seen in the voicing of stops such as $k > g$ and $t > d$ in some dialects, and the loss of occlusion as in $g > y$, $j > y$ and $p > v$ for example, while in other cases the stops are preserved, as in $g > g$ and $t > t$ in many cases. Overall, this presents as a partial change towards lenition and elision of the stops, with only minor variations between the eastern, western and northwestern dialects. The main variations are seen in the merger of the sibilants $\{ś, ṣ, s\} > s$, the nasals $\{ñ, ṇ, n\} > n$, and the liquids $\{r, l\} > l$ in the eastern dialect, while these are all retained in north-western dialect, and only the sibilants are merged in the western dialect²⁹¹.
- In the first and second century Gandhāran Buddhist manuscripts lenition of the intervocalic stops as both fricatives and approximants is recorded. While elision is also recorded in some cases, as in $\{c, j\} > y / \emptyset$, $\{p, b\} > v / \emptyset$, suggesting a more advanced

²⁹¹ E.g., Mehendale 1948: xxiii; 210-216 [Chapter IV, Synoptic Tables]; Salomon 1998b: 73-75.

stage of lenition than in the Aśokan inscriptions, the record of attested reflexes is mixed. The overall pattern suggests a partial weakening rather than elision of the stops, involving an intermediate stage of MIA development (§ 8.2.1.2.6.3).

- In the later stage literary Prakrits, as represented by Śaurasenī, the unvoiced stops originally became voiced, and apart from the retroflex series, all stops were subsequently weakened or elided, as in $T > \{y, v\} / \emptyset$ ²⁹², representing an ongoing trend in the weakening of these elements.
- In Pali, the OIA intervocalic stops are typically preserved, as are the semi-vowels, the phonemic status of \tilde{n} and n is retained although that of η is lost, and the sibilants merged²⁹³. Hence Pali shows a mixed development, including archaic features such as the retained stops and semi-vowels, but also some advanced features such as merger of the sibilants like the eastern dialect.

9.2.2 Consonant clusters

The OIA consonant clusters developed in similar ways among the MIA languages, but with variations occurring on both a diachronic and geographic basis. Changes typically involve both manner and place assimilation between consonants, with manner assimilation controlled by the stronger consonant in both regressive and progressive directions.

9.2.2.1 Comparative development: Aśokan inscriptions

A tabulation of individual cluster reflexes recorded in the northwestern, western and eastern Aśokan inscriptions is included in Table C2 in Appendix C, and an overview of the typical changes is presented in Table 9.2.2.1. Developments are described primarily in terms of the dominant tendencies for manner assimilation, although secondary palatalisation triggered by y and r is also common, and regional differences in place assimilations are also seen, such as $kś > kkh / cch$ in the eastern and western inscriptions respectively. Alternative forms are sometimes used in the same inscription, and along with orthographic errors²⁹⁴ it is difficult to gain a clear sense of the overall language development in some cases. While there are many similarities in the overall patterns of change, the assimilation of clusters appears to be more progressed in the east and west than in the northwest. Some of the main patterns of development include the following:

²⁹² E.g., Misra 1967: 162-172; Bubenik 2003: 206-207.

²⁹³ E.g., Oberlies 2019: 114-115, 131-132

²⁹⁴ E.g., Hultzsich 1925: lxxxvii-lxxxviii.

Table 9.2.2.1: Typical MIA reflexes⁽¹⁾ of OIA medial consonant clusters in the Aśokan Prakrits, expressed in terms of manner assimilation⁽²⁾

	Stop (T)	Nasal (N)	Fricative (S)	<i>l</i>	<i>v</i>	<i>y</i>	<i>r</i>
T +	NW: T ₂ T ₂	TT & others	TT / (S)		T _v / T	TT/T _i y	Tr
	E: T ₂ T ₂	TT & others	TT / (S)	Tl / TT	T _v	TT/T _i y	TT
	W: T ₂ T ₂	TT & others	TT / (S)	Tl	T _v / T _p	TT	TT/Tr
N +	NW: N _α T _α > ṁT _α	none				NN / (Ny)	
	E: N _α T _α > ṁT _α	none				NN / (Ny)	
	W: N _α T _α > ṁT _α	none				NN / (Ny)	
S +	NW: TT(h) / ST	SS/Sp				SS/(Sy)	Sr / SS
	E: TT(h)	mixed		Sv		SS/(Sy)	SS
	W: TT(h)	mh		Sv		SS/(Sy)	Sr / SS
l +	NW: lp > lp					ll	
	E: pp					l / y	
	W: pp					l / y	
v +	NW:					vv / vy	vr
	E:					vy	vv
	W:					vy / vv	vv
y +							
r +	NW: rT (Tr) / TT	NN	r S / SS		rv / vv	ry / yy	
	E: TT	NN	r S / SS		rv / vv	yy	
	W: TT	NN	r S / SS		rv / vv	ry / yy	

NW = northwestern (Gandhāra = shaded cells), E = eastern, W = western Aśokan Prakrit
Notes: (1) Only the MIA reflexes are shown
(2) Concurrent place assimilation also occurs in many cases but is omitted for clarity.

Source: compiled from Hultzsch (1925: lviii-lxii, lxxxvii-xc, cii-civ) and Mehendale (1948: xxiv-xxv, 21-27, 217-237 [Chapter IV, Synoptic Tables])

- Clusters formed between pairs of stops always develop through regressive place assimilation to form geminates of the second stop. Homorganic nasal + stop clusters remain stable and stop + nasal cluster tend to develop through progressive manner assimilation in favour of the stop, although other developments occur in some individual cases. The consistency of these changes across all regions suggests that they had progressed during an earlier period in the MIA language development.

- Ty clusters mostly develop as geminates, although epenthetic vowel insertions are recorded in a number of cases suggesting a level of retention (§ 10.1). The resolution of clusters through epenthetic insertions appears to be more common in the east.
- Tr clusters appear to be mostly retained in the northwest, partially retained in the west, and mostly assimilated in the east.
- Ny clusters often develop through palatalisation and assimilation as *ññ* geminates.
- ST and ST^h clusters are mostly assimilated in favour of the stop, although some *st* clusters are retained in the northwest and partially in the west.
- Sy clusters are either resolved through insertion of an epenthetic *i* vowel, as in Sy > *Siy*, or assimilated as geminate sibilants.
- Sr clusters are either retained or assimilated as geminate sibilants. The assimilation appears to have been more developed in the east, while retained clusters are more evident in the west and northwest.
- ly clusters assimilate as *ll* in the northwest, and as either *ll* or *yy* in the east and west. *vy* clusters are often retained, although assimilation as *vv* is also seen in the northwest and west. *vr* clusters are retained in the northwest, while assimilating as *vv* elsewhere.
- The *rC* clusters all develop in slightly different ways:
 - *rT* clusters are typically assimilated in the east and west, although a mixed development is seen in the northwest where they are often retained, although sometimes recorded orthographically as *Tr* (Hultzschi 1925: lxxxvii).
 - *rN* clusters are typically assimilated in favour of the nasal.
 - *rS* clusters show mixed development in all locations, being either retained or assimilated as a sibilant geminate.
 - The *r* + approximant *rv* and *ry* clusters are either retained or assimilated in favour of the second consonant to give *vv* and *yy* geminates.

9.2.2.2 Comparative development: Gāndhārī and later stage Prakrits

The changes recorded in the Gāndhārī Buddhist manuscripts, and the later period MIA literary Prakrit Śaurasenī, compared with OIA, are listed comparatively in Table 9.2.2.2. The Pali assimilation patterns (Oberlies 2019: 150-156) are essentially the same as the progressed MIA developments. The main points of comparison are as follows:

Table 9.2.2.2: Comparative development of Gāndhārī (G) and MIA Śaurasenī (Ś)⁽¹⁾
medial consonant clusters (MIA cases are in the shaded rows)

	Stop (T)	Nasal (N)	Fricative (S)	<i>l</i>	<i>v</i>	<i>y</i>	<i>r</i>
Stop (T) +	G: T ₁ T ₂ > T ₂ T ₂	<i>tm</i> > <i>tv</i> (<i>tt</i> , <i>pp</i>) <i>jñ</i> > <i>ññ</i>	TS > TS stable	Tl > Tl Tv > Tv		Ty > TT ⁽²⁾ Tr > Tr	
	Ś: T ₁ T ₂ > T ₂ T ₂	TN > TT <i>jñ</i> > <i>ññ</i>	TS > TT ^{h(2)(3)}	T{l, v} > TT (<i>dv</i> > <i>bb</i>)		Ty > TT ⁽¹⁾ Tr > TT	
Nasal (N) +	N _α T _α > N _α T _α stable	<i>nm</i> > <i>nm</i> small sample	<i>ṃS</i> > <i>ṃS</i> small sample	Nv > Nv small sample		Ny > NN ⁽²⁾ Nr > Nr small sample	
	N _α T _α > N _α T _α stable	N ₁ N ₂ > N ₂ N ₂	<i>ṃS</i> > <i>ṃs</i> coalesced	Nv > NN		Ny > <i>ññ</i> ⁽²⁾ Nr > NN	
Fricative (S) + ⁽³⁾	ST > ST / TT ST ^h > ST ^h /TT ^h mixed	SN > SN / SS uncertain <i>Sm</i> > <i>Sp</i> <i>hm</i> > <i>mm</i> / <i>m^h</i>	none in G	Sl > Sl rare <i>Sv</i> > <i>Sp</i>		Sy > SS ⁽²⁾ <i>Sr</i> > <i>Sr</i> / SS mixed	
	ST / ST ^h > TT ^{h(1)}	<i>Sm</i> > <i>mh</i>	SS uncommon	S{l, v} > SS		S{y, r} > SS	
<i>l</i> +	<i>l</i> T > TT	small sample	none in G			<i>ly</i> > <i>ll</i>	
	<i>l</i> T > TT	<i>l</i> N > NN	<i>l</i> S > SS		<i>lv</i> > <i>ll</i>	<i>ly</i> > <i>ll</i>	
<i>v</i> +		none in G				<i>vy</i> > <i>vv</i> <i>vr</i> > <i>vr</i> / <i>vv</i>	
		<i>v</i> N > NN				<i>v</i> {y, r} > <i>bb</i> (<i>vv</i>)	
<i>y</i> +							
<i>r</i> +	<i>r</i> T > TT / <i>r</i> T mixed	<i>r</i> N > NN	<i>r</i> S > SS / <i>r</i> S mixed	<i>rl</i> > <i>ll</i> <i>rv</i> > <i>rv</i> / <i>vv</i>		<i>ry</i> > <i>ry</i> / <i>yy</i> uncertain	
	<i>r</i> T > TT / <i>r</i> T	<i>r</i> N > NN	<i>r</i> S > SS	<i>rv</i> > <i>bb</i> (<i>vv</i>)		<i>ry</i> > <i>yy</i>	
<p>Changes are presented in the format of Skt. cluster > G or MIA reflex</p> <p>Notes: (1) The Pali assimilation is essentially the same as later Prakrits such as Śaurasenī</p> <p>(2) Includes palatalisation of coronals.</p> <p>(3) <i>ps</i> > <i>cch</i> also palatalised</p>							

Source: compiled from Table 8.3.6.2 and Misra (1967: pp 136-156)

- Similar developments are seen for clusters formed between pairs of stops or stops and nasals, and these are typically unchanged from those seen in the Aśokan inscriptions. These changes appear to have occurred prior to the Aśokan era and been stable at least until the time of the early literary Prakrits.
- All other consonant clusters developed fully as geminate clusters in Śaurasenī, through processes of regressive and progressive manner assimilation in favour of the stronger consonants, similar to the Pali developments (Oberlies 2019: 147-148, 150-152). This is an advanced stage of development compared with the eastern Aśokan Prakrit, where some of these assimilation patterns had not developed fully.
- Similar manner-based assimilation is seen in some of the Gāndhārī clusters, but to a more limited extent. This occurred fully in all the {T, N, S}y and rN clusters, but only partially in the Sr, r{T, S} and rv clusters, although some of the mixed reflexes may relate to borrowings or conservative orthography (e.g., § 8.3.4.4.2). These patterns are similar to the partial changes recorded in the north-western Aśokan Prakrit, suggesting that the language remained relatively stable over the intervening period.
- At the same time, a number of clusters had not assimilated or only assimilated to a limited extent in Gāndhārī, including the TS, Tl, Tv, Nv, Sl, Sv and ry clusters.

The only OIA clusters which develop fully through progressive assimilation of the weaker consonant in Gāndhārī are the Cy clusters. As discussed previously (§ 8.3.1.4.3), y is one of the weakest and most susceptible of the semi-vowels to lenition and assimilation due to its more open and vowel like articulation, and the Gāndhārī reflexes may reflect a partial stage progression of this type of MIA development.

9.2.3 Comparative overview: medial consonants and clusters

Some of the distinctive features of the OIA medial single and cluster consonants in Pali, the Aśokan Prakrits, Gāndhārī and the literary Prakrits, as described above, are summarised in Table 9.2.3. These features reflect both geographic diversity and diachronic developments, derived from source materials dating from the third century BCE to around the middle of the first millennium CE. Some of the important features include:

- Single intervocalic stops: these remain unchanged in Pali and are partially changed in the Aśokan Prakrits, while showing an intermediate stage of lenition in Gāndhārī, and a more advanced stage of lenition including deletion in the literary Prakrits.

Table 9.2.3: Comparative development of intervocalic single and cluster consonants in MIA

		Stops (T)	Medial clusters VC₁C₂V	Sibilants (S)	Semi-vowels (R)	non-labial Nasals (N)
Pali Various stages from 3 rd c BCE		Typically retained	Assimilated to geminates or aspirated geminates of the stronger element	{ś, ṣ, s} > s	y, r, l, v retained	ñ retained {n, n} > n
Aśokan Prakrits 3 rd c BCE	Eastern (Central similar)	Partially weakened – mixed reflexes	Mostly assimilated to geminates of the stronger element. Resolution by epenthesis quite frequent	{ś, ṣ, s} > s	y, l, v retained. r > l	{ñ, ṇ, n} > n
	Western	Partially weakened – mixed reflexes	Many assimilated, but those involving y and r often retained	{ś, ṣ, s} > s	y, r, l, v retained	ñ, ṇ, n retained
	North-western	Partially weakened – mixed reflexes	Many assimilated, but those involving y and r often retained	ś, ṣ, s retained	y, r, l, v retained	ñ, ṇ, n retained but ṇ < > n interchangeable
Gāndhārī (§ 8.2, § 8.3) 1 st & 2 nd c CE		Partially weakened towards S or R, as for example in <i>gh</i> > <i>g</i> / <i>h</i> and <i>d</i> > <i>ḍ</i> / <i>y</i>	Many assimilated, but those involving y and r often retained	ś, ṣ, s retained	y, r, l, v retained	ñ, ṇ, n retained but ṇ < > n interchangeable
Śaur. Prakrit ≈ 5 th c CE +		Mostly weakened or elided, as for example in <i>Th</i> > <i>h</i> and <i>T</i> > <i>y</i> , <i>v</i> / <i>Ø</i>	Mostly assimilated to geminates of the stronger element	{ś, ṣ, s} > s	y, r, l, v retained	{n, n} > n

- Medial consonant clusters: clusters are typically assimilated as geminates of the stronger element in the later literary Prakrits and Pali. This pattern did not develop fully in all clusters in the eastern and western Aśokan Prakrits, and was less advanced in the northwestern Prakrit and Gandhāran Buddhist manuscripts.
- Sibilant merger: the three OIA sibilants typically merged, mostly to the dental *s*²⁹⁵, in all of the languages except in the northwest where they are retained in both the Aśokan inscriptions and Buddhist inscriptions and manuscripts.
- Semi-vowels: although sometimes weakened or elided in intervocalic situations, the semivowels are typically retained except for the eastern Prakrit where *r* merged with *l*.
- Non-labial nasals: the most distinctive feature in the full merger of *ñ* and *ṇ* with *n* is the eastern Aśokan Prakrit and the literary Prakrits, while the distinction is retained in the western Prakrit. While all nasals are recorded in the north-western Prakrit and Buddhist manuscripts, it is likely that *ṇ* and *n* had merged while *ñ* only had a minor or restored phonemic status (§ 8.2.2), similar to the situation in Pali (Oberlies 2019: 115).

Many of these changes reflect the typical diachronic developments that occurred more widely in the MIA languages²⁹⁶, seen for example in the gradual weakening of the intervocalic stops and spread of cluster assimilation between the third century BCE Aśokan Prakrits, through the first and second century CE Gandhāran Buddhist manuscripts, to the later era literary Prakrits. In this context the development of the OIA medial consonants in Gāndhārī represents a relatively intermediate stage of MIA development²⁹⁷, similar to that seen in the word initial clusters (§ 9.1). At the same time retention of the three sibilants in Gāndhārī reflects a conservative pattern compared with the other Prakrits in which these had merged. Such conservatism may be related to the relative geographic isolation of Gandhāra and possible influences from the strong Sanskrit culture that existed in the northwest (§ 2.2.3), although this does not appear to have influenced other features such as the partial lenition of single intervocalic stops which became more advanced between the time of the Aśokan inscriptions and the later Buddhist manuscripts (§ 8.2). This suggests such local influences were not strong and may have only served to slow the rate at which some of the normal MIA changes developed in Gāndhārī. The changes recorded in the Aśokan inscriptions had most likely occurred over

²⁹⁵ These also appear as the palatal sibilant *ś* in Māgadhī Prakrit (Bubenik 2003: 207).

²⁹⁶ E.g., Misra (1967: 100-172) provides a comprehensive overview of these developments.

²⁹⁷ E.g., Salomon 1999: 124-125.

several centuries prior to the third century BCE, and been well developed by the middle of the first millennium BCE, prior to any scholarly engagement between later Buddhist communities and the Sanskrit grammarians in the northwest.

The intermixing of conservative and advanced features in Pali gives rise to some interesting patterns in the language, such as length contrasts between short and long stops. Although not unusual among the world's languages (§ 4.6), the ways in which this feature may have developed in Pali are unclear. For example, the lenition of medial stops is common on a cross linguistic basis (Campbell 2013: 37), and the changes in Gāndhārī and other Prakrits appear to reflect a natural development in this way, but one which appears to have been constrained in Pali. On the other hand, while the advanced assimilation of consonant clusters is reflected in the potentially quite early canonical language of Pali, this is less well developed in Gāndhārī and the other Prakrits of this time. This type of strength-based manner assimilation does not appear to be well documented on a cross-linguistic basis (§ 6.1.3), and linguistic motivations for such assimilation are analysed further in § 9.3, while possible reasons for the unusual contrasting developments in Pali are discussed in § 9.3.5.

9.3 Review of MIA manner assimilation in medial clusters

9.3.1 Strength based accounts

Aside from some combinations of the obstruent stop and nasal consonants, the predominant development in the OIA medial clusters in MIA and Pali is to form geminate clusters in favour of the more occlusive consonant, involving both manner and place assimilation in most cases (Oberlies 2019: 150-151, 154-156). Similar developments have occurred in Gāndhārī, but to a lesser extent and with less consistency, involving full assimilation of the Cy clusters only, and partial assimilation of some of the Cr and rC clusters (§§ 8.3.6, 9.1.2). The Pali assimilation patterns have been explained in terms of a consonant strength hierarchy, whereby stronger consonants assimilate weaker ones, regardless of directionality (§ 6.1.3). The relative strengths of the consonants are defined in terms of the degree of oral stricture, which is cast in the order of stops > nasals > (palatal stops) > sibilants > *l* > *v* > *y* > *r*, or T N S R *l v y r*²⁹⁸, although it is unclear whether the ordering of the semi-vowels has a clear phonetic basis, or is assigned

²⁹⁸ Oberlies 2019: 151; Suzuki 2002a: 97-98; 2002b: 64.

according to the observed patterns of assimilation (§ 6.1.3). The liquids *l* and *r* normally have a brief but partial contact during their articulation, and in this sense have greater stricture than the more open and vowel like glides. It has also been suggested that *r* may have been pronounced with a more open articulation as an alveolar continuant in OIA (§ 8.3.1.4.4), and the presence of *vr-* in OIA word initial onset clusters supports this analysis.

While this stricture-based approach offers a consistent account of the typological patterns observed in MIA medial clusters, manner features are not observed to spread consistently on a cross-linguistic basis. This can be seen for example in the converse patterns of cluster assimilation recorded in some of the regional European and Italian dialects, in which the more sonorant elements dominate (§ 6.1.3). And while one motivation for the MIA strength-based assimilation may have been to maximise the contrast between clusters and adjacent vowels (Suzuki 2002b: 64), such a development is not universal since the weakened clusters in the European dialects reduce such contrasts. Dissimilation between consonants having the same manner is also reported in some languages, as for example in *okto* > *oxto* and *ktena* > *xtena* in Modern Greek in which one of the stops is weakened (Gordon 2016: 142), and in /fθ/ > [ft] and /sθ/ > [st] seen in some varieties of English in which one of the fricatives is strengthened (Fromkin et al. 2021: 83). In these examples dissimilation appears to have been preferred to the formation of geminates through place assimilation, highlighting the difficulties in seeking a consistent account of manner assimilation in consonant clusters.

The partial stages of development recorded in Gāndhārī are interesting in this regard, since they involve assimilation of *y* and *r* which are classified as weakest in the Pali strength hierarchy, suggesting a tendency for this pattern to develop first in clusters with the most significant strength contrast. The semi-vowels have strong internal formants which provide reliable perceptual cues to their manner and place of articulation, and these lend a level of stability to clusters involving stops, nasals and fricatives with semi-vowels²⁹⁹. This is also reflected in the stability of these clusters in the OIA languages over a considerable period, and it is possible the less extensive changes to these clusters in Gāndhārī reflect a more natural and slower paced linguistic development in Gandhāra,

²⁹⁹ The nature of these perceptual cues is discussed in Appendix B. Refer also to the discussions in § 8.3.1.4, § 8.3.2.4, § 8.3.3.4 and § 8.3.4.4. The perceptual clarity of the semi-vowels in these clusters will also depend upon how distinctive their formant patterns are compared those of adjacent vowels, as for example for *y* and *i* which are very similar in this regard.

where Gāndhārī was more isolated from the levels of cross fertilisation that occurred between other Prakrits on the more open Gangetic plain. It is also interesting that while all of the {T, N, S}y clusters assimilated as geminates in Gāndhārī, the {T, N}r clusters are typically retained, while the Sr clusters are only partially assimilated. This suggests that y may be weaker than r in an assumed strength hierarchy of the semi-vowels, which differs from the Pali hierarchy of y > r. Although ry clusters are common in Sanskrit, there are no yr clusters, and the assumption that y is stronger than r is based on examples of regressive assimilation only, as in Skt. *ārya-* > P *ayya-* for example (Oberlies 2019: 150), which may be the result of anticipatory as much as strength-based assimilation (§ 8.3.4.4.4). In this context the Gāndhārī developments offer support for an alternative ranking of r > y in any assumed semi-vowel strength hierarchy.

The stability of the OIA TS clusters, and variable record of the ST clusters in Gāndhārī (§§ 8.3.1.3, 8.3.3.1), point towards the resilience of sibilants in clusters with stops, despite the relative strength relationship, and sibilants do not appear to assimilate easily to stops, even in a regressive position. This has been related to the strong perceptual cues of fricatives, and the ease of articulatory transition in these clusters (Appendix A4).

9.3.2 A composite approach

As discussed in § 4.6, patterns of Sanskrit consonant ‘doubling’ are described in ancient grammatical treatises, although these processes are largely omitted in contemporary grammars. It has been suggested there may be a diachronic link between these Sanskrit gemination patterns and MIA manner assimilation, and that the record of Sanskrit gemination may reflect a ‘change in progress’ (Murray 1982; Cho 1999: 152-185; Suzuki 2012, 2015). This has been analysed as a composite phenomenon involving several factors, which include lengthening of the oral closure, lengthening of the first consonant, and fortition of the syllable initial position (Suzuki 2012, 2015), while prosodic factors may also play a role (Kobayashi 2004: 32-34). The Sanskrit gemination patterns are reviewed in § 9.3.3, while the suggestion that MIA assimilation is a possible diachronic extension of these patterns is explored further in § 9.3.4.

9.3.3 Sanskrit gemination patterns

The patterns of Sanskrit gemination recorded by the ancient grammarians are presented in Table 9.3.3, which is a copy of the tabulation published by Suzuki (2015: 3-8,

Table 1), which was developed based on accounts of gemination presented by a number of Sanskrit scholars such as Whitney (1924: 78-79, §§ 227-229). It can be seen that there is a mostly direct correlation between those Sanskrit clusters in which the second consonant typically doubles, and those which experience regressive assimilation in MIA, as recorded in Table 9.2.2.2, and similarly between those clusters in which the first consonant doubles, and those where progressive assimilation occurs³⁰⁰. These groups of clusters are shaded and unshaded respectively in Table 9.3.3.

Table 9.3.3: Gemination of biconsonantal clusters C₁C₂ in Sanskrit

$\begin{smallmatrix} C_2 \\ C_1 \end{smallmatrix}$	T	N	S	<i>l</i>	<i>v</i>	<i>y</i>	<i>r</i>	<i>h</i>
T	T ₁ T ₁ T ₂	TTN	TTS/(TSS)	?TT <i>l</i>	TT <i>v</i>	TT <i>y</i>	TT <i>r</i>	—
N	NT	?N ₁ N ₁ N ₂	—	?NN <i>l</i>	?NN <i>v</i>	?NN <i>y</i>	?NN <i>r</i>	—
S	STT/(SST)	SNN/(SSN)	SS	?SS <i>l</i>	SS <i>v</i>	SS <i>y</i>	?SS <i>r</i>	—
<i>l</i>	<i>l</i> TT/(<i>l</i> IT)	? <i>l</i> lN	—	<i>ll</i>	? <i>l</i> <i>l</i> <i>v</i>	? <i>l</i> <i>l</i> <i>y</i>	—	? <i>l</i> <i>l</i> <i>h</i>
<i>v</i>	—	<i>v</i> NN/(<i>v</i> <i>v</i> N)	—	—	<i>vv</i>	<i>vy</i>	? <i>v</i> <i>v</i> <i>r</i>	—
<i>y</i>	—	—	—	—	—	<i>yy</i>	—	—
<i>r</i>	<i>r</i> TT	<i>r</i> NN	<i>r</i> S	? <i>r</i> ll	<i>rv</i>	<i>ry</i>	—	<i>rh</i> /(<i>rh</i> <i>h</i>)
<i>h</i>	—	<i>h</i> NN	—	? <i>h</i> ll	<i>hv</i>	? <i>h</i> <i>y</i>	<i>hr</i> /(<i>h</i> <i>h</i> <i>r</i>)	—
<i>ḥ</i>	<i>ḥ</i> TT	—	? <i>ḥ</i> S	—	—	—	—	—

? indicates that these examples are not reported in all of the source documents.
/() indicates less common variations.

Source: copied from the summary tabulation published by Suzuki (2015: 8, Table 1)

One interpretation of Sanskrit doubling suggests that it records a phonetic rather than a phonological process, in which lengthened consonant allophones are recorded using available phonemes (§ 4.6). In this understanding Sanskrit doubling is a record of phonetic tendencies, which are a precursor to the future phonological developments that occurred in MIA assimilation. Hence for example in Sanskrit *śakya*, which would be recorded as *śakka* in the doubling account, there is no contrast between -*ky*- and -*kky*-. This might also be represented as *śa^kkyā*, which reflects the possible allophonic nature of the development and uncertainties about the extent of lengthening. While there is some divergence of opinion as to how this process may have developed, and the extent to which

³⁰⁰ This section reviews the basis for manner assimilation between consonants having different manners. Stop + stop and nasal + nasal clusters, where the consonants have the same manner, assimilate through regressive place assimilation (§ 8.3.6.2), and are not considered in this analysis.

it contributes among other factors, including the role of prosodic factors (Suzuki 2015: 13-15), at a minimum level the record of Sanskrit gemination appears to have some association with the developments that occurred in MIA assimilation.

9.3.4 An analytical approach: articulatory characteristics and syllable patterns

Apart from clusters formed between stops and nasals, which typically maintain or develop place assimilation between these occlusive consonants, other OIA clusters fall into groups that have either an increasing or a decreasing strength profile. The patterns of MIA manner assimilation are examined in the following sections in the context of these groupings, while factors such as the robustness of segments based on articulatory characteristics and perceptual cues, and the syllable patterns in the language, are also considered.

9.3.4.1 Clusters having an increasing strength profile

These clusters correspond with the Sanskrit gemination patterns found in the shaded left-hand area of Table 9.3.3. In all these cases, the second consonant doubles in Sanskrit gemination, seen in the examples of $ST / SN > STT / SNN$ and $hC > hCC$ involving a prior fricative, and $lT > lTT / (llT)$, $vN > vNN$ and $rC > rCC$ involving prior semi-vowels. The Sanskrit clusters have falling sonority and typically divide as coda and onset across the syllable boundary, and although some of the ST and SN clusters occur word initially in OIA, these appear to develop as hetero-syllabic clusters in medial positions (Appendix A4). The presence of a coda consonant in the preceding syllable provides a prosodic weight to this syllable which satisfies any metrical requirements. There are several factors which may influence the development of manner assimilation in these clusters:

- Cluster final consonants typically exhibit robust external perceptual cues through formant transitions and release bursts to the following vowels, which are stronger than those of cluster initial consonants (§ 6.1.2). This appears to lend a greater stability to the second consonant, especially in the case of stops and nasals, even though initial consonants such as fricatives and semi-vowels possess strong internal cues. Due to the falling sonority and increasing occlusion of these clusters, the sequence of articulation is that of an extended closure prior to opening and release in the following vowel, as for example in $V > C_1 > . C_2 < V$, where $>$ and $<$ indicate closing and opening

respectively in this sequence, and the most occlusive consonant is highlighted in bold. This pattern suggest that C₁ is more likely to be weakened as the central element in a more gradual transition, compared with C₂ which is released in a single more precise transition.

- The relative perceptual weakness of the syllable final C₁ consonant was recognised by the Indian grammarians, who described this as incomplete articulation (Murray 1982: 172-174). Kobayashi (2004: 36-37) observes that since *r* cannot end a word in OIA, some ancient grammarians preferred that it should not end a syllable, leading to the insertion of *svarabhakti* after *r* in some clusters. Furthermore, geminates were preferred across syllable boundaries compared with clusters having different apertures, and doubling may have been motivated in order to create the same aperture across the syllable boundary (Kobayashi 2004: 31-34).

This suggests that regressive lengthening of the second consonant across the syllable boundary may have occurred as a phonetic response to the weakness of the more sonorant syllable final consonants in the prior syllable, thereby strengthening the syllable boundary. This might be represented for example as Skt. *sar.pa-* (MW s.v. m. “snake”) > *sar^p.pa-*, where the superscripted ^p indicates an allophonic lengthening. In this approach Sanskrit gemination represents an early stage in the regressive assimilation by *r* of the more strongly positioned *p*, which may have developed diachronically as *sar^p.pa-* > *sa^rp.pa-* > MIA *sap.pa-*, which also involves lenition and deletion of *r* during the oral closure to the newly formed geminate. The falling sonority of the syllable final *-r^p* > *-r^p* pair in the transitional stages is acceptable within normal syllable constraints, and coda sequences such as liquid plus obstruent or fricative plus stop occur in languages such as Burushaski and Finnish for example (Gordon 2016: 98).

9.3.4.2 Clusters having a decreasing strength profile

These clusters correspond with the Sanskrit gemination patterns found in the unshaded right-hand area of Table 9.3.3. In all of these cases the first consonant doubles in Sanskrit gemination, seen in the examples of TS > TTS involving a following fricative, and Tl > ?TTl, Tv > TTV, Sv > SSv, Ty > TTy, Sy > SSy, Tr > TTr and Sr > ?SSr involving following approximants. These clusters follow an opening sequence following an initial closure into the cluster from the preceding vowel, as in V > C₁ < C₂ < V, where

> and < indicate closing and opening respectively in these sequences, and the most occlusive consonant is highlighted in bold. This pattern suggests that C₂ is more likely to be weakened as the central element in a more gradual transition, compared with C₁ which has a clearer definition in the transition from the preceding vowel. About 50% of the OIA clusters having an increasing sonority profile also occur word initially and can therefore occur as medial onsets, while the remaining 50% do not occur word initially and divide as coda and onset across the syllable boundary (Appendix A4). Prosodic factors can also play a role in the lengthening of these clusters, which can syllabify as C₁.C₁C₂ for metrical reasons (§ 4.5). Progressive manner assimilation occurs in all of the MIA cases, and these clusters are analysed below in relation to these alternate syllable structures:

- Medial onset clusters having a V.C₁C₂V syllable structure, as in *ta.kra-* (MW s.v. n. “buttermilk”) and *śa.kya* (MW s.v. adj. “able”) for example, where *kr-* and *ky-* also occur word initially in OIA. In these cases, gemination of the first consonant appears to represent a lengthening and regressive projection of C₁ across the syllable boundary. This may be due to the perceptual weakness of stops in preconsonantal position, where their release bursts are blocked by the second consonant and they become more reliant on the V₁C₁ formant transitions (Jun 2004: 61-62). This may lead to a phonetic strengthening of C₁ which extends across the syllable boundary, as in *ta.kra* > *ta^k.kra* and *śa.kya* > *śa^k.kya* for example, where the superscripted ^k indicates an allophonic lengthening, which can also be thought of as an articulatory anticipation of C₁. The eventual MIA assimilation could develop through a strengthening of the coda C₁ and deletion of C₂, which is more prone to lenition in this space between the stop and vowel, as in *ta^k.kra* > *tak.k'a* > MIA *tak.ka* and *śa^k.kya* > *śak.k'a* > MIA *sak.ka* for example. This might also be interpreted as persistence of the first consonant into the second one.
- Medial coda-onset clusters having a VC₁.C₂V syllable structure, as in Skt. *svap.na* (MW s.v. m. “sleep”) and *kal.ya* (MW s.v. adj. “healthy”) for example, where *[pn- and *[ly- do not occur initially. This syllabification patterns against the preference of the Syllable Contact Law (§ 4.6), in that sonority rises rather than falls across the syllable boundary. Murray (1982: 164-165) has observed that languages will attempt to adjust to the SCL on a diachronic basis. This occurs through epenthesis in some cases, or gemination in cases such as *kal.ya* > **kal.^lya* and *svap.na* > **svap.^pna* for example

(p. 175), where the superscripted ^l and ^p indicates an allophonic lengthening. In these cases, the phonetic lengthening is motivated by a balancing the syllable contact sonorities, and at the same time is suggestive of the eventual development of MIA assimilation. The main issue with this approach is that the disallowed onset clusters *[pn- and *[ly- are recreated, albeit weakly, on a transitional basis. This appears to be overcome by deletion of the weaker second consonant concurrently with strengthening of the onset through gemination (Murray 1982: 176-177), as in *kal.ya* > **kal.^lya* > *kal.ṽa* > MIA *kal.la* and *svap.na* > **svap.^pna* > *svap.pⁿa* > MIA *svap.pa*.

9.3.4.3 Summary

The suggested phonetic and phonological developments associated with Sanskrit doubling and MIA assimilation are summarised in the following table. Superscripted segments represent a lengthening of consonants as diachronically transitional allophones.

C ₁ : C ₂ strength	Syllable structure	Gemination OIA > proto-MIA	Gemination motivation	MIA Development
C ₁ < C ₂	VC ₁ .C ₂ V	<i>sar.pa</i> > <i>sar^p.pa</i>	anticipation & strengthen syllable boundary	> <i>sa^p.pa</i> > <i>sap.pa</i> ,
C ₁ > C ₂	V.C ₁ C ₂ V	<i>ta.kra</i> > <i>ta^k.kra</i>	anticipation & strengthen syllable boundary	> <i>tak.k^ra</i> > <i>tak.ka</i>
C ₁ > C ₂	VC ₁ .C ₂ V	<i>kal.ya</i> > * <i>kal.^lya</i>	syllable contact law	> <i>kal.ṽa</i> > <i>kal.la</i>

In these approaches, the primary motivations are related to either strengthening of the syllable boundary, whereby stronger elements are anticipated across the syllable boundary, or simplification of the syllable boundary in order to satisfy the syllable contact law. These phonetic tendencies give rise to allophones which were recognised in consonant doubling patterns by the Sanskrit grammarians, but were not, apart from in isolated instances, established as phonological developments in Classical Sanskrit. At the same time, it is suggested these tendencies for gemination acted as a possible precursor to the development of assimilation in the vernacular Prakrits.

9.3.5 Sociolinguistic considerations

Sociolinguistic approaches to understanding the grammatical simplifications of the MIA languages were discussed in § 2.3, and it is worthwhile to briefly review the attested changes within this framework. It was suggested that as socially dominant OIA speakers

migrated into South Asia, the MIA vernacular languages may have developed through the adoption of an OIA dialect as a *lingua franca* among the indigenous populations, and although these developments might also be explained in other ways, the changes that occurred are not inconsistent with this interpretation. These included breakdowns of the case morphology and erosion of verbal forms, along with the sound changes such as assimilation of the more complex OIA consonant clusters. These phonological changes involved both regressive place assimilation between dissimilar stops to form geminate stops, as well as strength-based assimilation between consonants of different manner to form geminate consonants. Such changes offer a simpler articulation of the clusters based on length compared with the articulatory transitions required between different stops or between other consonant combinations in clusters (§ 6.1).

As discussed in § 9.2.2.1 the assimilation of OIA stop + stop clusters as geminate stops was fully developed in all the MIA languages by the time of the Aśokan inscriptions in the 3rd century BCE. At the same time the manner assimilation of clusters had not progressed fully in all the clusters, although this was more advanced in the east while being less so in the northwest. Regressive place assimilation is very common among the world's languages, and the early development of these changes point towards this occurring more readily in language change than manner assimilation processes, which as discussed in § 6.1 are observed less frequently in the world's languages. Anticipatory place assimilation involves a more straightforward articulatory adjustment of place only compared with the seemingly more complex processes of the regressive and progressive manner assimilation discussed in § 9.3.4. These appear to involve the development of phonetic tendencies to strengthen or adjust the syllable boundaries or account for metrical requirements, involving adjustments to both the manner and place and of articulation in many cases. This understanding is supported by the partial progression of these changes seen at the time of the Aśokan inscriptions, suggesting they had only developed gradually the languages over time. The more advanced nature of the changes in the eastern inscriptions, compared to the northwestern inscriptions, is also suggestive of a wider interaction with indigenous speakers on the Gangetic plain than in the more isolated environs of the northwest.

These changes contrast with retention of the original clusters in the conservative Sanskrit culture, where these may have been consciously retained, at least in the

intellectual and scriptural language of that culture, in order to maintain a distinction from the vernacular languages. While the pattern of the MIA changes had diversified slightly by the time of the Aśokan inscriptions, none of these had developed to the extent that the clusters were all fully assimilated, as is the case for the language recorded in the later stage literary Prakrits as well as in the Pali canon. The other interesting aspect of these changes relates to the intervocalic consonants, which had been weakened or merged with the approximants, to different degrees, by the time of the Aśokan Prakrits, while they are retained in Pali. As discussed in § 9.2.3, the lenition of medial stops is common on a cross linguistic basis, and appears to reflect a common development among the Prakrits, and it is difficult to explain why this had not developed in a similar way in Pali.

The origin and nature of Pali has been the subject of much scholarly debate³⁰¹, including whether it is a natural language, or had been restored or even “Sanskritised” in some way during formation of the oral and written canon. While the exploration of these issues falls outside the scope of this thesis, and there are many other aspects of the language which need to be considered in such an evaluation, a comparison between the Gāndhārī and Pali phonologies offers some interesting insights in relation to this issue. The extensive record of Gāndhārī Buddhist manuscripts displays a widespread use of transitional as well as historic and advanced phonological forms, indicating that all of these may have been recognised in the language or its dialects at the time, revealing a pattern of variable phonetic and phonological developments in the Gāndhārī language³⁰². This appears to reflect the absence of a clearly prescribed phonology in what was a relatively young scholarly culture, and the scribes appear to have recorded the phonological variability of change in progress in a relatively faithful manner through the variable orthography, thereby creating a record of the diachronic changes that had occurred or were occurring in the language at the time. In this context, it is not unreasonable to suggest that the phonological variations revealed in Gāndhārī reflect a general pattern of language variability in the wider body of MIA languages³⁰³. This is not evident in Pali, however, where, despite some variations, a much clearer pattern of unweakened single consonants and fully assimilated clusters suggests that the phonology

³⁰¹ E.g., Geiger 1943: 1-7; Norman 1992; Cousins 2013; Karpik 2019a, 2019b; Levman 2019; Oberlies 2019.

³⁰² The variations which arise between dialects found in central or high contact areas of language use, and those found in remote or insular areas, are well documented (Trudgill 2011: 3-9).

³⁰³ Similar variability is seen for example in the medial cluster reflexes for the eastern and western Aśokan inscriptions, which are tabulated in Table C2 (Appendix C).

was possibly rationalised and standardised in some way. It is possible this was done in order to clarify ambiguities in spelling and meaning resulting from variable forms, which most likely existed in early oral transmissions of the canon in the vernacular languages. And while the retention or restoration of unweakened single consonants might be interpreted as a form of “Sanskritisation,” this is not the case for the clusters however, in contrast with the restoration of some of the original OIA clusters seen in the third century CE Gāndhārī manuscripts (§ 8.7). It appears more likely that the Pali clusters may have been merged or rationalised into the most advanced forms, thereby avoiding the ambiguous historic and transitional forms such as those which are evident in Gāndhārī for example. Hence if adjustments were made to the Pali consonant phonology, these were not restorations towards Sanskritic forms, so much as a clarification or rationalisation of consonants into a clearer C and \bar{C} phonological system, while maximum modulation between vowels and consonants is retained and developed in a VCV and $V\bar{C}V$ type of structure. This both restrained the weakening of single medial consonants, as well as advancing the patterns of cluster assimilation, both of which were progressing in the MIA vernacular languages at that time. Whether Pali originated as a natural regional language, an early *lingua franca*, or even from the language spoken by the Buddha during his teaching life, this analysis, which must also be considered within the wider body of linguistic evidence, suggests that the Pali phonology was rationalised to some extent during the development and recording of the scriptural canon. The loss of geminate consonants in Ancient Greek from around the first century CE, and the ongoing development of the MIA languages³⁰⁴, highlights the conservatism and stability of Pali, which has been the formal language of Buddhist scriptures of the Theravāda communities for nearly two millennia.

³⁰⁴ E.g., Trudgill 2011: 4-5; Misra 1967: 195-198.

10. Common types of sound change recorded in Gāndhārī consonants

The following common categories of sound change have been observed on a reasonably regular basis in the Gāndhārī source materials.

10.1 Epenthesis

The insertion of epenthetic or *svarabhakti* vowels in the resolution of OIA consonant clusters is not uncommon in Gāndhārī (Allon 2001: 98), and this phenomenon has been recorded in many of the source materials, including in clusters involving stops, nasals, fricatives and semivowels. The cases recorded in § 8.3 are collated in Table 10.1, and where alternative reflexes are also recorded for particular clusters, these are included to indicate the nature of the variant reflexes, while the cells are also highlighted for clarity.

Table 10.1: Gāndhārī consonant cluster reflexes involving epenthesis

OIA Cluster	Gāndhārī reflexes	Source document and examples: Sanskrit > Gāndhārī For source references refer to Table 8.1 Alt: alternate reflexes. (#): atypical reflexes		Era
-dm-	-dum-	BLS4-1	<i>padmaḥ</i> > <i>paduma</i>	1 CE
	-tum-	BLS9-1	<i>padmī</i> > <i>patumaṃ</i>	
-dhm-	-dhum-	BLS1-1	<i>uddhmātakam</i> > <i>udh[u]maï</i>	1 CE
-pn-	-pun-	Aś-H	<i>prāpnoti</i> > <i>prapuṇati</i>	3 BCE
-kl-	-kil-	RS-12	<i>klāmatha-</i> > <i>kilamaṣa-</i>	2 CE
	Alt: <i>kr-</i>	RS-5	<i>klomakam</i> > <i>[kr]umao</i> complex etymology	
-kl-	-kil-	BLS9-1	<i>upakleśān</i> > <i>uvakileśa</i>	1 CE
	Alt: <i>-kr-</i>	BLS4-1	<i>śukla-</i> / <i>śukra-</i> > <i>śukra-</i> (Baums 2009: 676)	
gl-	gil-	BLS2-1	<i>glānakaḥ</i> > <i>gīlanago</i>	1 CE
	ghil-	BLS1-1	<i>glānaḥ</i> > <i>[ghil](*)a)no</i>	
kv-	kuv-	BLS2-2	<i>kvacit</i> > <i>kuvaci</i>	1 CE
	Alt: <i>kv-</i>	BLS1-1	<i>*kvathya-</i> > <i>kvachia-</i>	
dv-	duv-	Aś-H	<i>dvau-</i> > <i>duv[i]-</i>	3 BCE
		Aś-M	<i>dvi-</i> > <i>duvi-</i> , <i>duve</i> (Mānsehra)	
		BLS2-2	<i>dve-</i> > <i>[du]ve-</i>	1 CE
	Alt: <i>du-</i>	BLS2-2	<i>dviprakārā-</i> > <i>dupragara-</i>	1 CE
		BCS5-1	<i>dve-</i> > <i>[du]e-</i>	

-ty- -ṇdy-	Alt: dv-	RS-12	<i>dvītyam</i> > <i>d(*u)die</i>	2 CE
		RS-12	<i>dvādaśa-</i> > <i>dvaḍaśa-</i>	2 CE
		RS-20	<i>dvāraiḥ</i> > <i>dvarehi</i> , (<i>ṣaḍ-dvāram</i> > <i>ṣa-dvaro</i>)	
	-tiy-	InsK-M	<i>caitya-</i> > <i>cetiya-</i> , <i>pratyamśa-</i> > <i>pratiyamśa-</i> (#)	3 CE
	-ṇdiy-	Aś-H	<i>Pāṇḍya-</i> > <i>Paṇḍiya-</i> (#)	3 BCE
	Alt: -nd-	Aś-H	<i>Pāṇḍy-a</i> > <i>Paṇḍa-</i> (#)	
-nv- -ṇy-	Alt: -ṇ-	After nasals (§ 8.3.2.4.2)		
		InsK-M	<i>*tanvaka-</i> > <i>tanuvaka-</i>	2 CE
		BLS4-1	<i>samanvāgata-</i> > <i>samuṇagada-</i>	1 CE
	-niy-	RS-20	<i>samanvāgataḥ</i> > <i>-samuṇakada-</i>	2 CE
		Aś-H	<i>ānṛiṇya-</i> > <i>anaṇiya-</i>	3 BCE
-hl- śl- -hy- -śy- sy- -sy-	-hal- śil- -hiy-	After fricatives (§§ 8.3.3.4.1, 8.3.3.4.3)		
		InsK-M	<i>Bāhlīka-</i> > <i>Bahalia-</i>	1 CE
		BLS2-1	<i>ślokena</i> > <i>śilogano</i>	1 CE
	Alt: (y)y	InsK-M	<i>mahyam</i> > <i>mahiya</i>	2 CE
		RS-19	<i>-uhyate</i> > <i>avuyadi</i>	2 CE
	-śiy-	Aś-H	<i>prativeśya-</i> > <i>prativeśiya-</i>	3 BCE
	Alt: ś(ś)	InsK-M	<i>Kāśyapīya-</i> > <i>Kaśavia</i>	1 CE
		Aś-H	<i>syāt</i> > <i>siyati</i>	3 BCE
	-siy-	InsK-M	<i>syāt</i> > <i>si(y)ati</i> > <i>siati</i>	1 CE
		BLS2-1	<i>syāt</i> > <i>siyadi</i>	
		BLS2-2	<i>syāt</i> > <i>siyadi</i>	
	Alt: s-	RS-12	<i>syāt</i> > <i>asa</i>	2 CE
		Aś-H	<i>rabhasye</i> > <i>rabhasiye</i>	3 BCE
		InsK-M	<i>kaniṣkasya-</i> > <i>kaniṣkasya-</i>	2 CE
	-s(s)-	BLS1-1	<i>ālasya-</i> > <i>[a]lasa-</i> and many similar	1 CE
vy- -rm- -rv-	vi(y)- -rum-	After semivowels (§§ 8.3.4.4.2, 8.3.4.4.4)		
		RS-20	<i>vyādhi-</i> > <i>viāsi-</i> . Similarity of <i>y</i> and <i>i</i> , as in <i>vya-</i> > <i>via-</i> , or <i>vya-</i> > <i>*viya-</i> > <i>via-</i> ?	2 CE
		InsK-M	<i>Buddhavarman-</i> > <i>Budhavaruma-</i> (#)	3 CE
	Alt: -rm-	Aś-H	<i>karma-</i> > <i>krama-</i> , <i>dharmā-</i> > <i>dhrama-</i> (#)	3 BCE
		InsK-M	in <i>dharmarava-</i>	1 BCE
	-ruv-	Aś-H	<i>pūrva-</i> > <i>pruva-</i> (#) (Possibly metathesis)	3 BCE
	Alt: -rv-	InsK-M	<i>sarva-</i> > <i>sarva-</i>	1 BCE

<i>-ry-</i>	<i>-riy-</i>	BLS1-1	<i>sarva-</i> > <i>sarva-</i>	1 CE
		Aś-H	<i>ānamtaryena</i> > <i>anamtariyeṇa</i> (Anaptyxis)	3 BCE
		Aś-M	<i>mādhurya-</i> > <i>madhuriya-</i>	3 BCE
	<i>-ri(y)-</i>	InsK-M	<i>ācārya-</i> > <i>ayaria-</i> (<i>-ry-</i> > <i>*riy-</i> > <i>-ri-</i>)	1 CE
		BLS1-1	<i>vīryavatām</i> > <i>viri[v](*)ataṇa</i> uncertain	1 CE
		BLS1-2	<i>vīrya-</i> > <i>(*)vi)ria-</i>	
		BLS2-2	<i>vīryendriyaḥ</i> > <i>viridrigo</i>	
		BCS5-1	<i>jāgaryā-</i> > <i>jagaria-</i>	2 CE
		RS-20	<i>ārya-</i> > <i>aria-</i>	
	<i>-rig-</i>	BLS2-2	<i>sūryam</i> > <i>surigo</i> (#)	1 CE
	Alt: <i>-ry-</i>	InsK-M	<i>ācārya-</i> > <i>acarya-</i> , <i>bhāryā-</i> > <i>bharya-</i>	1 CE
		BLS2-1	<i>kāryam</i> > <i>karya</i> among others	1 CE
	Alt: <i>(y)y</i>	Aś-H	<i>āryeṣu</i> > <i>-ayeṣu</i>	3 BCE

It has been suggested that cases of vowel epenthesis recorded in the Nid-G by the BL scribe 4 only occur in loanwords (Baums 2009: 155), reflecting borrowings from other MIA languages. While this pattern may be common, the extent to which epenthetic vowels are found in loanwords in other Buddhist manuscripts is typically not indicated in the source documents. At the same time comparative synoptic tables are available for the Aśokan inscriptions which facilitate an assessment in these cases. It has been observed that epenthesis is more common in the eastern dialect than the other Aśokan Prakrits (Salomon 1998b: 73-74) and since the Aśokan inscriptions were composed in this dialect these forms may have been copied as loanwords in the north-western and western inscriptions. The Aś-H and Aś-M examples listed in Table 10.1 include Skt. *prāpnoti* > G. *prapūṇati* (*-pn-* > *pun-*), *dvau* > *duv[i]*, *dvi-* > *duvi-*, *duve* (*dv-* > *duv-*), *prativeśya-* > *prativeśiya-* (*-śy-* > *-śiy-*), *rabhasye* > *rabhasiye* (*-sy-* > *-siy-*), *pūrva* > *pruva* (*-rv-* > *-ruv-*), *ānamtaryena* > *anamtariyeṇa* and *mādhurya-* > *madhuriya-* (*-ry-* > *-riy-*), most of which also appear in parallel or similar eastern and western forms³⁰⁵.

The main epenthetic vowels which are recorded are the palatal and labial vowels *i* and *u*, which typically occur prior to dental or palatal and labial consonants respectively, as in *kl* > *kil* or *dv* > *duv* for example. In these clusters the vowel qualities develop in anticipation of features of the second consonant, emerging in the space that occurs as the

³⁰⁵ E.g., Hultzsch 1925: 183-213, Appendix; Mehendale 1948: 217-237.

articulators move between the two consonants. The deployment of /i/ as an epenthetic vowel is quite common on a cross linguistic basis (Gordon 2016: 165-166), and is seen in the Gāndhārī examples of *kl* > *kil*, *gl* > *gil*, {*ś*, *s*}*y* > {*ś*, *s*}*iy*, and *ry* > *riy*. The OIA semivowel *l* is a dental lateral approximant and the articulatory movement of the tongue between the velar stops *k* or *g* and *l* involves transition through a high front position, which appears to trigger the emergence of epenthetic *i* in this case, while the other examples suggest the development of *i* in anticipation of the following palatal *y*. In the same way epenthetic *u* typically appears in anticipation of following labial consonants, seen in the Gāndhārī examples of *dm* > *dum*, *pn* > *pun*, *kv* > *kuv*, *dv* > *duv*, *nv* > *nuv*, and *rm* > *rum*. The singular case of *-hl-* > *-hal-* appears to involve the insertion of epenthetic schwa [ə], which is also common on a cross linguistic basis (Gordon 2016: 165). It is possible that in this case there isn't enough closure in the initial *h-* to allow the following consonant to condition a perceptible vowel quality. The number of alternative reflexes used by the scribes suggests a level of uncertainty in the recording of these sounds, whether as developing allophones in Gāndhārī or as borrowings from other dialects.

Epenthesis can be motivated by prosodic as well as phonetic factors, and when epenthetic vowels are inserted to satisfy syllabic, stress, or metrical requirements they are often recognised within the phonology, while epenthetic sounds which arise as phonetic or allophonic elements may remain unrecorded (Gordon 2016: 161-168). Since all of the listed MIA and Gāndhārī cases emerge in conditioned environments, and other reflexes are recorded for many of these clusters, it seems likely they have a phonetic origin. This supports the observations that the Gāndhārī scribes were attempting to record the perceived sounds of the language, providing further insight into the patterns of language change in progress at the time.

10.2 Palatalisation

Palatalisation is common on a cross linguistic basis (Campbell 2000: 35-36), and is observed in the development of a number of reflexes in Gāndhārī, commonly associated with the presence of the palatal consonants *c*, *j* and *y* in original OIA clusters. These include the following cases, which typically involve regressive spread of the palatal place feature to cluster initial consonants:

- Dental stop + y clusters: these develop in initial positions as *ty-* > *c-* and *dhy-* > *j-*, involving palatalisation of the stop and concurrent deletion of the *y*, while in medial positions *-ty-* > *-c(c)-*, *-thy-* > *-(c)ch-*, *-dy-* > *-j(j)-* and *-dhy-* > *-j(j)-* indicate palatalisation of the stop along with concurrent assimilation of the *y* (§ 8.3.1.4.3).
- Nasal + y clusters: the development of *{-ṇ-, -n-}y-* > *-ñ(ñ)-* / *-ṇñ-* indicates palatalisation of the retroflex and dental nasals along with concurrent progressive manner assimilation of *y* by the nasal, to form a geminate palatal nasal (§ 8.3.2.4.3).
- Fricative + y clusters: the development of *-ṣy-* > *-ś(ś)-* indicates palatalisation of the sibilant along with concurrent assimilation of the semi-vowel. This development is not seen in the case of *-sy-* > *-s(s)-* in which the dental sibilant is retained. The developments of *-hy-* > *{-ś-, -z-, -j-, -(y)y-}* all involve palatalisation, while also suggesting an unclear phonetic realisation (§ 8.3.3.4.3).

The following examples are also observed in three consonant clusters:

- The development of *ṁ.jñ* > *(ṁ).ñ* suggests that *anusvāra* reflects the palatal nasal, as in *ñ.ñ* (§ 8.3.5.2).
- The development of *r.ty* > *(c).c* in *martyaka-* > *macaga-* by the BL scribe 4 indicates palatalisation of the geminate stop, while the development of *r.thy* > *r.th* in *nisāmarthya-* > *ṇisanartha-* by the BC scribe 5 may reflect a conservative form in which the palatalisation is not recorded (§ 8.3.5.4).

10.3 Deaspiration

Deaspiration of intervocalic voiced aspirate stops is recorded in a number of Gāndhārī reflexes, such as *-gh-* > *-g-*, *-dh-* > *-ḍ-* for example, also involving lenition in the latter case (§ 8.2.1.2.1, § 8.2.1.2.4). Deaspiration is also recorded in the development of some clusters, as in *-ggh-* > *-(g)g-* in *taggha-* > *taga-* (RS-12), and *-dgh-* > *-(g)g-* in *udghāṭana-* > *ugaḍaṇa-* (BLS4-1) (§ 8.3.1.1). These both involve an OIA cluster of the type *-Cgh-*, and although they are the only recorded reflexes of such clusters, it is possible they are following the general trend for deaspiration of voiced aspirates. The production of voiced aspirates requires more articulatory effort than required for unvoiced aspirates, suggesting lenition and articulatory ease as a motivating factor in these changes.

10.4 Aspiration

Sibilant + stop clusters: aspiration develops in typical reflexes of the OIA word initial *st-* and *sp-* clusters involving the dental and labial stops, seen in *st-* > *th-* and *sp-* > *ph-*, involving spread of the fricative place features to the second member of the cluster as the glottal aspirate [h] in the post-stop position. A similar development is seen in the word medial reflexes of the retroflex cluster *-ṣṭ-* > *-(t)ṭh-*, while the medial dental cluster is typically retained apart from the one example of *-st-* > *-(t)th-* seen in *śāstā* > *śathu* by the BL scribe 2. Although the velar case is typically recorded as *sk-* > *k-* without aspiration, there is some evidence that recording of the aspirated stop *kh-* may have been intended in some other instances (§ 8.3.3.1). Development of the glottal fricative cluster *-hv-* > *-(b)bh-* reflects both the occlusion of *v* and subsequent metathesis of the glottal fricative as aspiration (§ 8.3.3.4.2).

Atypical cases are recorded in *dg* > *(g)gh* in *-pudgalam* > *-pughalu* and *udga-* > *ugha-* (BLS1-1), *kt* > *(t)th* in Skt. **viṣakti-* > *G visathi-* (BLS4-1), and *bd* > *(d)dh* in *śabdam* > *śadha* (RS-12). The writing of *gh* for *g* is a common graphic habit of BL scribe 1, while *visathi-* appears to be an orthographic variant, and *śadha-* is counter to the dialectic trend for the loss of distinction between aspirates and non-aspirates (§ 8.3.1.1).

10.5 Retroflexion³⁰⁶

Although the OIA dental clusters are typically retained or give rise to dental geminate reflexes in Gāndhārī, some examples of retroflexion are seen in the cases of *-rt-* > *(t)ṭ-* and *-rth-* > *-(t)ṭh-* (§ 8.3.4.4), involving the progressive spread of features to the second member of the cluster. In these examples the retroflexion is retained even though the conditioning element is assimilated.

10.6 Metathesis

Two types of metatheses are commonly observed in Gāndhārī reflexes:

- The aspiration of sibilant + stop clusters, as in *ST* > *(T)Th* (§ 8.3.3.1), involving metathesis and spreading of the fricative features as aspiration of the stop. This type of

³⁰⁶ This is sometimes referred to as ‘cerebralisation’ after the older term ‘cerebral’ which has also been used to describe retroflex consonants in Sanskrit.

metathesis often occurs in sequences of sibilant + stop or stop + sibilant, where there can be an auditory difficulty in assessing the order of fricative noise compared with other acoustic cues in the train of speech (Gordon 2016: 172).

- Metathesis of the liquid *r* is well attested in the Gāndhārī source materials, and involves a reversal of ordering with both preceding vowels as well as following consonants. This is recorded most persistently in the Aśokan inscriptions, seen for example in Skt. *kirti-* > G *kitri-*, *karma-* > *krama-* and *dharmā-* > *dhrama-*, *darśana-* > *draśana-*, *-darśin-* > *-drasi-* and *pārśamda-* > *praśamda-*, *sarva-* > *savra-* and *pūrva-* > *pruva-*, involving a variety of *r* + C clusters containing stops, nasals, sibilants and other semi-vowels (§ 8.3.4.4.3). Other examples recorded in the manuscript sources include Skt. *sudurbalam* > G *so[dhro]ba(*lo)* (Lenz 2003: 45), *nirvidā* > *ṇivriḍa-*, *parvatam* > *[pra]vaḍa-* (Glass 2007: 124), *durgatim* > *dhroghadi* (Salomon 2008: 130), *dīrgha-* > *driga-*, *duḥśīlo* > *drośilo*, *pravrajanti* > *parvayaḍi* (Lee 2009: 59), and *dīrgha-* > *driga-* (Marino 2017: 133). Examples of anticipatory and intrusive *r* are also reported, as in Skt *gandharvaḥ* > G *ghadhrarvo*, *-pariṇāman* > *-prariṇamu*, *gandham* > *ghadro* (Allon 2001: 98), and *dharmam* > *dhrrama* (Marino 2017: 133). Metathesis of the liquid *l* is also seen in *pāṭaliputre* > *palaḍipu(*tr)e* (Lenz 2003:133). In examples such as Skt. *-vṛttih* > G *-vurti* (Salomon 2000: 91) and *ṛddhi-* > *i[dh]ri-* (Salomon 2008: 343), vocalic *r* develops as consonantal *r* through inclusion of an epenthetic vowel.

Metathesis of liquid *r* is an areal feature of both the ancient and modern northwest languages, and has also been referred to as “Dardic metathesis”³⁰⁷. The auditory cues of liquids are potentially extended, which may confuse a listener’s perception of their place in the word, and this is also described as perceptual metathesis (Gordon 2016: 170).

Metathesis can occur as a regular change, but mostly appears to occur on a sporadic basis (Campbell 2013: 33; Gordon 2016: 168). It has been suggested that metathesis can be analysed as a form of perceptual optimisation, whereby elements are gradually moved to locations where their perceptual cues are optimised, although such analysis is not universally agreed (Gordon 2016: 172-173).

³⁰⁷ E.g., Salomon 2000: 91; Allon 2001: 97; Lenz 2003: 45.

10.7 Occlusion and loss of nasality in sibilant + labial clusters

Reflexes of OIA medial clusters involving sibilants + labial consonants in many cases develop as a combination of sibilant + labial stop, as in $-S\{p, m, v\} \rightarrow -Sp-$, involving loss of nasality in the case of the Sm clusters and occlusion in the case of Sv (§§ 8.3.3.1, 8.3.3.2, 8.3.3.4.2). There is some variation in development of the sibilants, which are all palatalised in the Sp clusters, developing as $-\{\acute{s}, \text{ṣ}, s\}p \rightarrow -\acute{s}p-$, while their place is typically retained in the Sm clusters, developing as $-\{\acute{s}, \text{ṣ}, s\}m \rightarrow -\{\acute{s}, \text{ṣ}, s\}p-$. Similar patterns are also seen in the word initial clusters, with the place of the sibilant typically retained. It is suggested that closure into a stop from the controlled fricative articulation offers an easier articulation than transitioning to a partial closure of the lips in the labio-dental approximant v , while the loss of nasality in $Sm \rightarrow Sp$ also provides a slightly simpler articulation (§ 8.3.3.4.2).

11. Conclusions

The main aim of this thesis has been to document the consonant phonology of the Gāndhārī language. Gāndhārī is a MIA language which was spoken in the Gandhāra region of modern-day northern Pakistan and eastern Afghanistan, known through inscriptional and manuscript records dating from around the third century BCE until the fourth century CE. This work has been carried out through collation and analysis of the phonology recorded in published descriptions of Gāndhārī source materials by modern scholars. These materials include a substantial body of Buddhist manuscripts which became available during the past thirty or so years, along with the Aśokan rock inscriptions and other shorter Gāndhārī inscriptions accessed by western scholars in the late nineteenth and early twentieth centuries. All of the Gāndhārī source materials are written in the Kharoṣṭhī script, which was adapted from the Aramaic script to record the MIA Gāndhārī language. Changes to the phonology have been examined in comparison with the OIA languages, known today through the record of Vedic and Classical Sanskrit, and comparisons also made with other documented MIA languages including Pali. The Gāndhārī reflexes of OIA consonants and consonant clusters are often recorded using variant graphemic forms. These include historic, transitional, and advanced forms, which appear to reflect variations in the pronunciation of the language such as dialectic variations, phonetic remnants of historic language variation, or natural allophones in the spoken language. While these variant reflexes appear to reflect the absence of a clearly prescribed phonology in what was a relatively young scholarly culture, the scribes appear to have recorded the phonological variability of change in progress in a relatively faithful manner through the variable orthography. This record of variable forms offers fascinating insights into patterns of language change in progress, and ways in which the variant reflexes can be categorised and represented have been proposed. Another interesting feature of the Gāndhārī phonology is the partial assimilation of OIA consonant clusters, compared with the more complete assimilation which occurred in Pali and other MIA languages, involving strength based manner assimilation of clusters. The transitional nature of these changes in Gāndhārī offers insights into these processes of change, highlighting progressive stages in these developments, particularly with respect to clusters involving semi-vowels. Ancient Gandhāra lies adjacent to the northwestern mountain regions of South Asia, which are home to the modern-day Dardic languages. There appear to be some shared areal features between Gāndhārī and this group of languages, and the

Gāndhārī phonology has been compared with that of Torwali, a Dardic language located in the upper Swat valley area of ancient Gandhāra. Sociolinguistic approaches to understanding the overall simplification of the MIA languages compared to the OIA ancestor languages have also been explored. More detailed descriptions of these various features of the language studied in this thesis are provided in the following sections.

11.1 Collation and analysis of the Gāndhārī phonological record

The phonological record of the consonants and consonant clusters in the Gāndhārī language presented in the selected source materials has been collated into a tabular format, and organised systematically on the basis of the manner and place of the consonants and clusters. Word level examples from each of the source documents, which demonstrate the changes that occurred between the OIA and Gāndhārī consonants, in both word initial and medial positions, form the basis of these tabulations. A number of aspects of these changes have been examined, including the variant forms used to represent the Gāndhārī phonemes in the Kharoṣṭhī orthography, and the changes have been reviewed in the context of linguistic parameters and broader cross linguistic typologies. Interpretations by previous scholars have been reviewed where appropriate, and an overall assessment of each of the changes is presented in a shorthand format. Detailed summaries of the results of this work are included at the end of each particular section, providing a consolidated record of all of the phonological changes which occurred between the OIA and Gāndhārī languages, based on the published source materials available at the time of writing this thesis. Interpretation of the actual nature of the phonemes represented in the script, and appropriate forms of phonemic representation, have also been considered. An overview of some of the common types of sound change recorded in the Gāndhārī materials is also presented, including epenthesis, palatalisation, deaspiration, aspiration, retroflexion, metathesis and occlusion.

11.2 Categorisation of the variable phonemic record in a shorthand format

One of the problems associated with the record of variable phonological forms in a language like Gāndhārī is to determine which speech sounds are being represented by the orthography, where variant reflexes are used inconsistently in many instances. A system has been developed where all of the Gāndhārī reflexes recorded for each OIA consonant or consonant cluster are classified and annotated in a shorthand format according to whether

they are historic, transitional, advanced, borrowed or in some cases alternative forms. A concise record of these patterns is provided through this system, which has proven useful in the analysis of different aspects of the language. While the assigned classifications are open to interpretation, it is believed that the basic format of this system provides a useful tool for the ongoing interpretation and analysis of the language phonology.

11.3 Changes to OIA single consonants in Gāndhārī

11.3.1 Summary of distinctive phonological features of Gāndhārī

The following general patterns of development of the single consonants are observed in the source materials, as summarised in § 8.2.6.1:

- Word initial consonants generally remain stable
- Word medial voiceless stops merge with their voiced counterparts
- A general tendency for lenition of the intervocalic voiced stops to fricatives or approximants, representing change in progress. Evidence suggests that deletion was not typical, and most stops were subject to phonetic weakening while retaining their phonemic status as stops.
- The OIA nasals *m* and *n* are retained, the distinction between *ṇ* and *n* was lost, and *ṇ* only arises in geminate reflexes of OIA consonant clusters.
- The three OIA sibilants *ś*, *ṣ* and *s* are retained.
- The OIA semi-vowels *y*, *r*, *l* and *v* are typically retained, although *y* tends to be elided in the environment of palatal vowels.
- Word final consonants are typically lost.

These patterns compare with the general developments in Pali where the medial stops are retained unweakened, while they are weakened as approximants or elided in later stage MIA Prakrits. In both these cases the three OIA sibilants are merged as the dental *s* (§ 9.2.3).

11.3.2 Interpretation of phonemes and phonemic representation

Understanding the intended phonological and phonetic targets represented by the variant forms of reflexes, and the best approach to representing them in the lemma forms

of words, is another aspect of the Gāndhārī phonology which has been addressed in this thesis. This mainly relates to the intervocalic stops and some of the nasal consonants, as discussed in § 8.2.1.2.6 and § 8.2.2.2.2. Although particular approaches to this issue have been suggested, these are more by way of addressing the topic rather than presenting definitive answers. The graphemic variants recorded in the Gāndhārī reflexes most likely represent actual variations in pronunciation, and it is evident that all of the intervocalic stops were weakened in the spoken language to some degree, and pronounced as either fricatives or approximants. In the case of the OIA velar and dental stops, *k* and *g*, and *t* and *d*, these are mostly represented in Gāndhārī using either their historic forms, or transitional forms such as /-g-/ and /-d-/ for example. These reflexes are assumed to indicate a fricative pronunciation, and effectively represent allophonic variants in this particular environment. While the use of forms such as /-g-/ and /-d-/ could be helpful in indicating this weakened pronunciation, it is suggested that the original graphemes are retained to represent the phonemes in the lemma forms of words in these cases. At the same time, it is reasonable to recognise the well-established voicing mergers of (*t*, *d*) > *d*, and (*k*, *g*) > *g*.

In other cases, however, the pronunciation of intervocalic stops appears to have been close to or merged with that of other phonemes, and the graphemes of these phonemes were used as reflexes in some of the sources. This is particularly so in the cases of *b* > *v*, *j* > *y*, *bh* > *h*, *dh* > *s*, and *gh* > *h* for example. These reflexes are, however, intermixed with more conservative forms, and it is unclear whether a full sound merger had actually occurred, similar to the later stage MIA Prakrits, or whether the use of these graphemes was an orthographic convenience based on a similar, but not identical pronunciation. For this reason, it is considered appropriate to retain reference to the original OIA phonemes, while at the same time there could be value in marking the intervocalic stops with an underscore diacritic in transliteration, such as /-b-/, /-j-/, /-bh-/, /-dh-/ and /-gh-/ for example. This marking, which borrows and extends from the system used by the RS scribe, provides a reference to the significant phonetic weakening that had occurred, whether as an allophone or merger with another phoneme. None of these adjustments involve changes to the phonemic inventory of the language, but are rather modifications to the representation of these graphemes to reflect the extent of transitional sound changes occurring in particular environments, although the full extent of these changes remains uncertain.

The graphemes for the dental and retroflex nasals, *n* and *ṇ*, have been merged orthographically by most of the scribes, who typically used one of these graphemes to represent both nasals. This suggests that these phonemes were not contrastive in Gāndhārī, and while most of the scribes followed a convention of writing *ṇ*, it is suggested that the dental nasal *n* is a more appropriate underlying form and representation for these nasals.

11.3.3 The Gāndhārī phonological system

A chart of the Gāndhārī phonological system has been presented in Table 8.2.6.2. This includes separate representations for both the word initial and word medial consonants, while some of the variant forms recorded for medial consonants are also listed.

11.4 Changes to OIA consonant clusters in Gāndhārī

11.4.1 Summary of distinctive features of consonant clusters in Gāndhārī

The following general patterns of development of the consonant clusters are observed in the source materials. While these represent the typical developments, variations are also recorded in many cases, which are discussed in more detail in § 8.3.6:

- Apart from the weakening of some of the OIA clusters to a single consonant, word initial clusters generally remain stable.
- Word medial clusters between occlusive stop + stop and nasal + nasal pairs are generally assimilated as geminates of the second member of the pair through regressive place assimilation. Stop + nasal clusters undergo mixed developments, while homorganic nasal + stop clusters remain stable.
- The OIA medial clusters having rising sonority typically remain stable, apart from the progressive assimilation of *y* in the *-{T, N, S, l, v}y-* clusters to form geminates of the first element.
- The OIA medial clusters of the ST and ST^h type show mixed results, remaining stable in some cases, and assimilating as geminates of the stop in others.
- Other OIA medial clusters having falling sonority, which are mostly of the *-r{[T, N, S, R]}*- type, show mixed results, remaining stable in some cases, and assimilating as geminates of the stronger consonant through regressive assimilation in other cases.

These patterns compare with the general developments in Pali and the later stage MIA Prakrits where all of the word initial clusters are reduced to single consonants, in most cases, while most of the medial clusters are assimilated as geminated of the strongest element (§ 9.2.3).

11.4.2 Manner assimilation in OIA consonant clusters

The manner assimilation of OIA consonant clusters as geminates of the strongest consonant in Pali and other MIA languages, including Gāndhārī, does not appear to be a common development on a cross-linguistic basis. While the changes in Pali have been explained in terms of a consonant strength profiles, this approach falls short of offering a consistent functional account. Other approaches to understanding these changes have been explored in § 9.3, including the possibility that they may reflect an extension of phonetic tendencies for gemination across syllable boundaries, as recorded by the ancient grammarians in the Sanskrit doubling accounts (§ 9.3.3). In this analysis gemination is explained as a strengthening of the syllable boundary in some cases, and a simplification of the syllable boundary in others, depending on the syllabic structure of the OIA clusters, while OIA metrical patterns may also contribute to the way clusters pattern across the syllable boundary and contribute to doubling effects. The partial extent of manner assimilation which occurred in the Gāndhārī clusters offers insights into these changes. The examples of manner assimilation in Gāndhārī mostly occur in clusters involving the *y* and *r* semi-vowels, which are classified as weakest in the Pali strength hierarchy, suggesting a tendency for this pattern to develop first in clusters with the most significant strength contrast. It is also interesting that while all the {T, N, S}*y* clusters assimilated as geminates in Gāndhārī, the {T, N}*r* clusters are mostly retained, while the *Sr* and *rC* clusters are only partially assimilated. This suggests that *y* may be weaker than *r* in an assumed strength hierarchy of the semi-vowels, which differs from the assumed Pali hierarchy where *y* is stronger than *r*. Although *ry* clusters are common in Sanskrit, there are no *yr* clusters, and the assumption that *y* is stronger than *r* is only based on examples of regressive assimilation, which may be the result of anticipatory as much as strength-based assimilation in Pali. In this context the Gāndhārī developments point towards an alternative ranking of *r* as stronger than *y* in any assumed semi-vowel strength hierarchy, offering important insights to the way these clusters may have developed in the MIA languages.

The semi-vowels have strong internal formants which provide reliable perceptual cues to their manner and place of articulation, which helps understand the relative stability of clusters involving stops, nasals and fricatives with semi-vowels. This is reflected for example in the stability of these clusters in the OIA languages over a considerable time period. The stability of the OIA TS clusters, and the variable record of the ST clusters in Gāndhārī, also point towards the resilience of sibilants in clusters with stops, despite the relative strength relationship, and sibilants do not appear to assimilate easily to stops, even in a regressive position. This has been related to the strong perceptual cues of fricatives, and the ease of the articulatory transition in these clusters. It seems reasonable to suggest that the less extensive changes in all of the Gāndhārī clusters reflect a slower paced linguistic development in Gandhāra, where the language was more isolated from the levels of cross fertilisation that the other Prakrits and Pali were exposed to on the Gangetic plain, and setting Gāndhārī apart from other peninsular developments in the MIA languages.

11.5 A record of language change in process

The variable representations of consonant reflexes in the Kharoṣṭhī inscriptional and manuscript materials, which include a range of historic, transitional and advanced forms, provide a fascinating insight into the processes of sound and language change which had been or were occurring in the two thousand year old Gāndhārī language. These variable forms have been categorised in a shorthand format developed for this thesis (§ 11.2), which facilitates understanding the patterns of language change in progress at that time. The recording of variable reflexes is common for both single consonants and consonant clusters, and the details have been summarised in various sections of this thesis, including the following:

- The weakening of single intervocalic consonants, as discussed in § 8.2.1.2.6.
- The variable recording of many of the cluster reflexes, as discussed in § 8.3.7.
- The variable recording of epenthetic vowel insertions, where alternate unchanged reflexes are also recorded in many cases, as discussed in § 10.1.

It is most likely that the graphemic variations reflect actual variation in the pronunciation of the language, and the pattern of variable reflexes indicates that the Gāndhārī scribes were attempting to capture an adequate orthographic record of the language within a continuum of language change, which included a divergent range of

conservative, transitional, allophonic, innovative and dialectic pronunciations. While these variant forms may also suggest the absence of a clearly prescribed phonology in what was a relatively young scholarly culture, the scribes appear to have recorded the phonological variability of change in progress in a relatively faithful manner through the variable orthography. The scribes were in effect trying to record a changing language in written form, for which the orthography was not fully developed, in an environment of variable pronunciation, while influenced by the orthography of the texts they had inherited from other language areas. And while it is difficult to discern any overarching hierarchy or common pattern of change among the variable reflexes, these make more sense when viewed as a record of language change in progress. Other changes such as the intermixing of orthographic *n* and *ṇ* (§ 8.2.2.2.2), the general breakdown or loss of the Sanskrit sandhi system (§ 8.6), and the transitional stages of cluster assimilation compared with later stage MIA Prakrits (§§ 8.3.7, 9.2.3), are also indicative of the general patterns of language change which were occurring at that time.

As discussed in § 1.4.4, when change occurs in a language from an archaic to an innovative form, this always involves a stage where both forms co-exist in the speech of individuals, as phonetic remnants of language variation remain embedded in the spoken language. And in written forms of language, it is found that non-standard orthographies are often phonetically based, while word alternations are often recorded in early scripts having an undeveloped orthography, and manuscripts may also incorporate many of the changes which occurred in a language's sound system prior to its script coming into existence³⁰⁸. In these contexts, it is evident that the variable phonological reflexes seen in the Gāndhārī inscriptional and manuscript materials provide an extensive record of patterns of language change in progress. As observed by Campbell (2013: 194), the study of this type of variation can make an important contribution to the understanding of linguistic change, and the record of change seen in the ancient Gāndhārī inscriptions and manuscripts is particularly valuable in this regard.

11.6 Comparison of developments in Gāndhārī and other MIA languages

The developments that occurred in Gāndhārī have been reviewed in comparison with those in the later stage MIA Prakrits, including Pali, and a number of distinct regional

³⁰⁸ E.g., Matsumoto 2019: 199; Schneider 2002: 80; Lahiri 2015; Gogoi, Morey and Pittayaporn 2020.

characteristics have been observed, which are consistent with those reported in earlier scholarship (§ 2.1). Some of the main differences include an intermediate stage of lenition of the intervocalic stops in Gāndhārī, compared with unweakened stops in Pali and full lenition of the stops in the Prakrits, partial assimilation of the consonant clusters in Gāndhārī, compared with near complete assimilation in Pali and the Prakrits, and retention of the three OIA sibilants in Gāndhārī compared with mergers which occurred in Pali and the Prakrits. A full analysis of these differences is presented in § 9.2.3 and summarised in Table 9.2.3.

Many of these changes reflect the typical diachronic developments that occurred more widely in the MIA languages, seen for example in the partial weakening of the intervocalic stops and spread of cluster assimilation in the first and second century CE Gāndhārī Buddhist manuscripts, compared with the full development of these changes in the later era literary Prakrits. In this context the development of the OIA consonants in Gāndhārī represents a relatively intermediate stage of MIA development. On the other hand, the changes recorded in the Aśokan inscriptions had most likely occurred over several centuries prior to the third century BCE, and may have been well developed by the middle of the first millennium BCE, prior to any scholarly engagement between later Buddhist communities and the Sanskrit grammarians. At the same time retention of the three sibilants in Gāndhārī reflects a conservative pattern compared with the other Prakrits in which these merged into either the dental sibilant *s* or the palatal sibilant *ś*. This may reflect shared areal influences between Gāndhārī and the northwestern Dardic languages, which also retain the three unvoiced sibilants (§ 11.8).

11.7 Scribal variations

The recording of the Gāndhārī language in the Kharoṣṭhī manuscripts occurred among a group of scholarly communities who diverged in location, time and possibly the dialect of members. And while there are some noticeable divergences in the graphic practices, such as in the use of diacritic marked reflexes by the RS and some of the BC scribes, but not the BL scribes, there is also much commonality of practice between the communities. This is evident in the common use of orthographic *ṇ* to represent both *n* and *ṇ* by most of the BL scribes as well as the RS and BC scribes, in the regular elision of the velar and palatal stops and *y* in the environment of palatal vowels by all of the scribes, in the interchangeable use of the same orthographic variants in Gāndhārī reflexes of many of

the Sanskrit consonant clusters, and in the regular use of common borrowed forms for technical terms from eastern dialects, for example. At the same time, alternations in the use of many of the graphic variants, both by individual scribes and within communities, suggests a degree of uncertainty or lack of agreement in the correct forms of representation, despite the evident desire to record the phonetic variations in a reasonably faithful manner. This can however be considered in the context of some of the broader linguistic influences on the ways in which Gāndhārī was recorded, including the incomplete development of the Kharoṣṭhī script, the more advanced developments seen in eastern source texts, wider areal influences from the northwestern Dardic languages and other languages passing along the trade routes passing through the region, along with the apparent absence of a clearly prescribed phonology and grammar in what was a relatively young scholarly culture, all shared among scholarly communities who diverged in location, time and possibly the dialect of members. And while the more conservative Sanskrit intellectual culture of the region was another possible influence on the Gāndhārī scribal communities, Sanskrit scriptures and knowledge were conveyed through an esoteric oral tradition at this time, and it is uncertain how a complex oral composition such as Pāṇini's grammar, for example, could have been shared widely among the Gāndhārī scribal communities. In these contexts, and despite the observed variations, the creation of a reasonably consistent set of orthographic practices in the recording of Gāndhārī speaks of a vibrant intellectual culture within the Gandhāran Buddhist community, but one which does not appear to have clearly resolved its understanding of the language and how best to represent it. The culture was evidently aware of the patterns of sound and language change that had occurred or were occurring between the older forms of the OIA languages and the more advanced MIA forms seen in eastern source texts, while at the same time attempting to capture an adequate functional record of the local Gāndhārī language and its dialects within this continuum of change.

11.8 Relationship to Dardic languages

The geographical proximity of ancient Gandhāra to the region of the Dardic languages in the northwestern mountain regions of South Asia suggests the likelihood of some form of affinity between Gāndhārī and this language group, and past scholarship has suggested that Gāndhārī may be most closely aligned with Torwali, a modern day language of the upper Swat valley. There is however a long gap in the diachronic record

between the end of the Gāndhārī manuscript tradition around the fourth century CE and the first recording of the Dardic languages during the nineteenth century, which makes it difficult to trace any direct historical connections between the languages. Compared with Gāndhārī the Dardic languages have developed more complex phonological systems, including in the case of Torwali classes of affricates and voiced sibilants, while both voiced and voiceless stops are retained medially. And although pronunciation of the stops is typically weakened in medial positions, they have retained phonemic status, in contrast with the MIA languages where they were eventually weakened and elided. On the other hand, and unlike the other MIA languages Gāndhārī has retained the three voiceless sibilants, a common feature among the Dardic languages, while rhotic metathesis, which is well attested in Gāndhārī, and especially in the Aśokan inscriptions, is a feature shared with many Dardic languages. And while the wider differences between Gandhari, Torwali and the other Dardic languages make it difficult to establish any clear historical connection between the languages, shared features such as the three sibilants and rhotic metathesis suggest the presence of some form of historical areal influence between them.

Another important consideration are the different geographical terrains and sociolinguistic environments in which Gāndhārī and the Dardic languages developed. Gāndhārī is a language of the lower valley plains and larger communities of Gandhāra, who were connected with the wider plains and MIA language communities of the Gangetic plain, while Torwali is a language of the upper Swat valley, spoken in a more remote high valley location, largely isolated from the changes occurring to the main Gāndhārī dialects. In this context it is expected that the languages might tend to develop in different ways, for the grammatical and phonological systems of Gāndhārī to simplify, and for those in Torwali and the other Dardic languages to remain stable or become more complex (§§ 2.3, 11.9). Shared areal features such as retention of the three sibilants and rhotic metathesis, especially in the Aśokan inscriptions, provide strong indications of an early areal connections between Aśokan or pre-Aśokan Gāndhārī and the northwestern languages. Given their close geographical proximity, it is not unreasonable to suggest that Torwali could have developed from an early Gāndhārī dialect, or that both languages shared a common ancestor language, not dissimilar to Aśokan or pre-Aśokan Gāndhārī. At the same time, it is expected they would have developed in quite different ways, each influenced by its own distinct and contrasting geographical and sociolinguistic environment.

11.9 Sociolinguistic considerations

The overall simplification that occurred to the vernacular MIA languages, compared with the OIA ancestor languages, potentially had its origins in sociolinguistic influences that arose during intermixing of the migrating OIA speakers with indigenous populations. This may have occurred both prior to and following the migrations of OIA speakers into South Asia. A sociolinguistic interpretation of these changes is supported by the simplifications that occurred in all aspects of the grammar, compared with the comparative stability of OIA Sanskrit used by the educated Brahmin of the northwest. In this interpretation, the vernacular MIA languages underwent significant changes due to sociolinguistic influences that arose in the more vibrant social interactions of the wider society, including between the IA speakers and indigenous populations, compared with the relatively isolated and self-contained culture of the socially dominant Brahmins. It is suggested that the MIA vernacular languages may have developed through the adoption of an OIA dialect as a lingua franca among the indigenous populations. Phonological changes such as the weakening and merger of stops with approximants, along with the gemination of consonant clusters, represent processes of simplifications of the phonology through the loss of contrasting forms. Such changes result from the diminished capacity of adults for grasping and pronouncing certain redundant phonetic and grammatical structures, especially when a more dominant contact language is adopted as a lingua franca by second language learners over a long period. Evidence of early change to the MIA languages is recorded in references to popular forms found in the early Vedic texts, and it is possible that the major changes occurred at a relatively early stage in the development of MIA. This understanding is also compatible with the early stage development of place assimilation in OIA stop + stop clusters to form geminate stops, and the slower paced partial development of manner assimilation between dissimilar consonants, as first recorded in the third century BCE Asokan inscriptions. While regressive place assimilation between dissimilar stops involves straightforward articulatory adjustments, processes of regressive and progressive manner assimilation appear to involve more complex adjustments, and proceeded at a slower pace of change. It is suggested that MIA gemination may have developed as an extension of phonetic tendencies to strengthen or clarify syllable boundaries, first recorded as Sanskrit doubling by the early grammarians. Although speculative, these approaches offer a way in which the MIA developments can be interpreted.

Comparisons between the Gāndhārī and Pali phonologies provide useful insights into questions about the nature of Pali. It is suggested that the record of variable reflexes in the Gāndhārī phonology reflects a similar pattern of uncertain and variable phonemic forms and pronunciation in the wider body of MIA languages. This compares with the Pali phonology which shows a distinct pattern of unweakened single consonants and fully assimilated geminate consonant clusters. This suggests that the Pali phonology may have been rationalised or adjusted in some way from its original vernacular forms during development and recording of the canon. In both cases these adjustments may have served to clarify ambiguous spellings arising from the presence of variable phonemic forms, through the deletion of uncertain transitional forms in both cases. Whether Pali originated as a natural regional language, an early lingua franca, or even from the language spoken by the Buddha, this analysis, which must naturally be considered within the wider body of linguistic evidence, supports the view that the Pali phonology was possibly rationalised to some extent during development and recording of the scriptural canon. Pali developed as a unique scriptural language having a clear phonological structure involving contrasts between short and long consonants, but distinct from both the OIA ancestor languages and the loosely defined evolving vernacular languages such as Gāndhārī and the other Prakrits.

11.10 Suggestions for further work

There is certainly more scope for ongoing research in many of the areas covered by this thesis, and areas which are likely to be of most interest include the following:

- A record of the links between the Gāndhārī consonant and cluster reflexes and all of the original OIA phonemes and clusters from which these derive, expressed as a reversed pattern of the diachronic phonological changes. While this might be undertaken as a fairly simple exercise based on the data reported in this thesis, a more thorough approach, in which the relative proportions of the variant forms of reflexes are collated in each case, will be more beneficial. This is a significant task in which these proportions must be counted in each of the published source documents where they are not already amalgamated. This material can be presented in tabulated or diagrammatic formats.
- A similar more detailed analysis of the orthographic habits of each of the scribes or monastic groups may also be beneficial, and this could be integrated during the preparation of a detailed reverse phonology.

- A similar level analysis of the vowel phonology, which can be integrated with the consonants in a more complete representation of the Gāndhārī phonology.
- Available knowledge of the Gāndhārī morphology can also be amalgamated and analysed in similar ways, especially in regards to the extent of weakening and levelling of the declensional case morphology.
- Although beyond the immediate scope of this thesis, the following are worthy of a more detailed exploration than could be afforded here:
 - further investigation of possible areal influences between Gāndhārī and the wider body of Dardic languages, including Aśokan and later stages of Gāndhārī.
 - the wider question of sociolinguistic influences on the grammatical simplification of the MIA language, including the morphology and verbal forms as well as the phonology.

Appendix A1: Retroflex consonants in OIA

Retroflex consonants were not present in the reconstructed PIE language (Clackson 2007: 34), and arose in a variety of assimilation conditions during the diachronic development of OIA from PIIr, as well as through the influence of borrowings from the native Dravidian languages at later stages³⁰⁸. These assimilated and borrowed phones attained phonemic status to varying degrees. While the diachronic developments between PIE, PIIr and OIA are complex, the earliest advance of retroflexion in OIA appears to have been the development of OIA *ṣ* from PIE *s and the PIIr palatal *š. From a synchronic perspective, the spreading of retroflexion between the OIA consonants is described in a series of well-known sandhi rules (§ 4.2), which typically describe the spread of retroflex place features into otherwise dental consonants. The following summary of these rules is based on the overview presented by Kobayashi (2004: 141-144):

- RUKI rule: $s > ṣ / \{r, ṛ, u, k, i\} _ \{\neq r, ṛ\}$: *s* becomes *ṣ* when preceded by *r*, *ṛ*, velar stops, and high vowels, while the spread is blocked by any following *r* or *ṛ*. Examples are *akṣara-* (MW s.v. n. “syllable”), and *tisṛ-* (f. “three”) where the spread is blocked. Both the semi-vowel *r* and vocalic *ṛ* are pronounced with an alveolar or palatal contact according to the ancient phonetic treatises (Allen 1957: 54-55), and feature as conditioning influences in several of the rules. In this case the tongue tip retains an alveolar or palatal position during pronunciation of the following sibilant.
- Retroflex *ṣ* spreads retroflexion to following dental stops: $T > Ṭ / ṣ _$. Examples are $\sqrt{duṣ} + ta > duṣṭa-$ (“spoiled”) and $\sqrt{tṛṣ} + ta > tṛṣṭa-$ (“thirsted”), in a common process of adjectival participle formation from a verbal root, but *das + ta > dasta-* (“wasted”) without the preceding high vowel. In cases such as $\sqrt{diś} + ta > diṣṭa-$ (“shown”) and $\sqrt{dṛś} + ta > dṛṣṭa-$ (“seen”) the palatal *ś* changes to *ṣ* before the dental stop (Macdonell 1927: 27), which is subsequently assimilated as *ṭ* in a further development.
- Retroflex stops spread retroflexion to following dental stops at morpheme but not at word boundaries, as for example in $\sqrt{id} + te > iṭṭe-$ (“praises”).
- Both *ṣ* as well as *r* and *ṛ* spread retroflexion to a following dental *n* provided this is not blocked by an intervening dental or palatal consonant, and *n* is followed by a sonorant. This is described in the so called NATI rule: $n > ṇ / \{r, ṛ, ṣ\} \{\neq C_{den}, C_{pal}\} \dots (+, \#) _$

³⁰⁸ Kobayashi 2004: 141-148; Fortson 2010: 211.

{V, n, m, y, v}. This spreading may transmit remotely, even across morpheme and word boundaries when joined by other sandhi rules. Examples are $\sqrt{tr}\text{ṣ} + nā > trṣṇā$ - (f. “thirst”) in the formation of a substantive from the verbal root, and *putreṇa* (m. “with the child”) in the instrumental singular ending *-ena* for *a*- nouns.

- *ṇ* is typically inert except that it retroflexes any following *n*, as in $\sqrt{k}\text{ṣud} + na > kṣuṇṇa$ - (adj. “trampled”) for example, involving several stages of sandhi, whereby *d* > *t* at the morpheme boundary, and *t* > *n* / *_n*, but *-nn-* > *-ṇṇ-* / *u_*. It is unclear whether *ṇ* spreads retroflexion to following dental stops, and the process of retroflexion in words such as *daṇḍa*- (m. “a stick”) and *piṇḍa*- (m. “a ball”) is uncertain (Kobayashi 2004: 141).

Diachronic patterns suggest that retroflexion was present and developing in OIA during the pre and early Vedic periods, while contact with the Dravidian languages appears to have reinforced and hastened the spread and phonemicisation of these sounds³⁰⁹. While the extent to which the contact influences overlapped with the developing internal changes is difficult to ascertain, it is possible that early contact lingua francas may have played a role in these developments. This complex history has resulted in an uneven spreading of retroflexion in OIA, and this reflects in the varying extent of phonemicity attained by each class of retroflex consonants (Kobayashi 2004: 142-143).

- The retroflex fricative *ṣ* appears to have complex origins stemming back as far as Proto-Indo-Iranian, and is also present in some lexical items in the Ṛgveda where its origins are unclear, as in *kavaṣa*- (MW s.v. m. “a shield”), *caṣāla*- (m. “a hive”) and *jalāṣa*- (n. “water”)³¹⁰ for example. Minimal pairs such as *asta*- (adj. “thrown”) and *aṣṭa*- (adj. “marked”) occur, and *ṣ* appears to have a clear phonemic status that is grounded in the early development of OIA (Misra 1967: 64-66; Kobayashi 2004: 143).
- While the retroflex stops *ṭ*, *ṭh*, *ḍ* and *ḍh* mostly appear as allophones of the dental stops when preceded by *ṣ*, they are also seen in a number of independent lexical items, such

³⁰⁹ E.g., Misra 1967: 64-65; Fortson 2010: 211.

³¹⁰ While **casāla*- and **jalāsa*- do not appear in comprehensive dictionaries such as Monier-Williams (1899), for example, *kavaṣa*- and *kavaṣa*- (m. “armour”) form a minimal pair providing support for the phonemic status of *ṣ*.

as *kaṭuka-* (adj. “sharp”), *kīkaṭa-* (adj. “poor”), *jaṭhala-* (“ocean”) and *śirimbiṭha-* (m. “a cloud”)³¹¹ for example, which are traced back to early PIIr (Kobayashi 2004: 143).

- Most occurrences of *ṇ* appear to be conditioned and it was for the most part an allophone of *n* in Vedic times. However, a few words with unconditioned *ṇ* occur, such as RV *pāṇi-* (m. “shop”), *maṇi-* (m. “ornament”), *guṇa-* (m. “strand”), *sthūṇa-* (n. “post”) and *sthāṇu-* (adj. “stationary”)³¹² (Misra 1967: 75; Kobayashi 2004: 143), while *n* and *ṇ* do contrast in some (but seemingly not many) minimal pairs, such as *anu-* (prep. “after, along”) and *aṇu-* (adj. “minute”), giving *ṇ* phonemic status at a minor level. It is suggested that *ṇ* became phonemic at a later stage in OIA due in the most part to borrowings from non-Aryan sources such as Dravidian (Misra 1967: 75; Fortson 2010: 211), which resulted in a seeming analogical spread of retroflexion into formerly dental nasals in Sanskrit words. One other situation in which retroflex *ṇ* arises is in homorganic clusters with stops of the same class, as for example in clusters such as *kaṇṭha-* (m. “neck”), *paṇḍita* (m. “learned man”), and *aṇḍa-* (n. “egg”). Similar to the NATI spread of retroflexion, this is another conditioned allophone of *n*.

While the presence of some minimal pairs confirms varying degrees of phonemicity among the retroflex consonants, the extent to which these occur as a proportion of the words involving retroflex consonants is difficult to determine without an extensive analysis of the lexicon. This would of necessity involve the tracing of word etymology in terms of diachronic and synchronic allophonic spread, along with the identification of loanwords and any subsequent analogical spreading.

The retroflex stops appear to have been pronounced with the tongue tip turned back against the palate³¹³. Allen (1953: 52-56) observes that descriptions of the tongue position given in the ancient texts can be imprecise, and descriptions of the related retroflex rhotics, the semi-vowel *r* and vocalic *ṛ*, are suggestive of an alveolar position. Ladefoged and Maddison (1996: 25-28) report that retroflex *ḍ* in some varieties of Hindi is apical, with the tip of the tongue placed against the alveolar ridge, while in Tamil this is sub-apical with the underside of the tongue placed in a post alveolar position (p. 27

³¹¹ **kaṭuka-*, **kīkaṭa-*, **jaṭhala-* and **śirimbiṭha-* do not appear in comprehensive dictionaries such as Monier-Williams (1899), for example, and do not form minimal pairs that support the phonemic status of *ṭ* and *ṭh*.

³¹² **pāṇi-*, **maṇi-*, **guṇa-*, **sthūṇa-* and **sthāṇu-* do not appear in comprehensive dictionaries such as Monier-Williams (1899), for example, and do not form minimal pairs that support the phonemic status of *ṇ*.

³¹³ E.g., Whitney 1924: 16 [§ 45]; Macdonell 1927: 16

[Figure 2.11]). This is similar to the description given in one of the ancient texts (Allen 1953: 53). The influence of Dravidian in the ongoing development of OIA retroflex consonants supports the idea that both pronunciations were possible at different times and regions. Languages tend to build consonant inventories in ways that maximise perceptual distinction while minimising articulatory effort, and only adopt more complex articulations once articulatory spaces requiring less effort have been exploited (Gordon 2017: 62-63). In a suggested three tier system, retroflex sounds are classified within a middle tier of articulatory difficulty, and do not represent an overly complex development.

Appendix A2: *Anusvāra*

Traditionally, *anusvāra* occurs in post-vocalic positions prior to consonants and is described as a ‘pure nasal’, generated by the post-vocalic breath passing through the nose. Although typically unmodified by any consonant, it was influenced by following consonants in Vedic, leading to nasalised glides, or possibly a nasalised transition between vowel and consonant articulation. When used correctly these ‘after sounds’ occur in two well-defined contexts³¹⁴:

- *Anusvāra* primarily occurred prior to fricatives including *h*, as in *vaṁśa-* (MW s.v. m. “cane”) and *siṁha-* (m. “a lion”) for example, and in some conditions prior to *r*.
- The usage was extended in Classical Sanskrit to replace Vedic nasalised \tilde{y} , \tilde{v} and \tilde{l} before *y*, *v* and *l* respectively.

The recording of *anusvāra* in these contexts suggests a narrow transcription of allophonic vowel nasality or after sound in certain environments, which may have been pronounced as labio-dental nasal [m] (Allen 1953: 20, 40). This might be expressed as [Ø] > [̃] / V_ {*s, ṣ, ś, h, (r), y, v, l*}.

In a further development *anusvāra* was used to mark word final nasals preceding a stop, usually *m* but also *n*, where homorganic nasals had been used previously. This was later extended to morpheme boundaries and even intra-morpheme positions within a word. This latter usage does not represent nasalisation in the same sense, but is a form of orthographic abbreviation that became quite widespread³¹⁵.

The traditional texts offer divergent descriptions of the phonemic nature of *anusvāra*, which vary between nasalisation of the preceding vowel, a nasal fricative or glide, a nasal consonant, a consonant or a vowel, which may reflect regional differences (Allen 1953: 41-45). While Whitney (1924: 24-25) suggested that *ṁ* represented a nasalisation and lengthening of the preceding vowel, as in *vāśa* for example, both Allen (1953: 41-45) and Cardona (2013: 1, 33) prefer the interpretation of *anusvāra* as a post-vocalic nasal segment, as in *vāśa* for example, although its exact phonetic and phonemic nature is difficult to define. Ladefoged and Maddieson (1996: 298-300) distinguish between oral vowels, contextually nasalised vowels which only have a perceptible nasality

³¹⁴ E.g., Macdonell 1916: 10; Allen 1953: 39-41.

³¹⁵ Macdonell 1916: 10; Whitney 1924: 70-73 [§§ 70-73]; Allen 1953: 40-46.

during the onset to following nasal consonants, and phonologically nasalised vowels having a much longer nasal airflow and perceptual nasality, and which contrast with oral vowels in many languages. While vowel nasalisation is uncommon in OIA (Allen 1953: 39-40), *anusvāra* appears to have originated as a perceptible nasalisation which occurred prior to sibilants and glides, for reasons that are not phonetically certain.

Misra (1967: 79-81) observes that *anusvāra* always occurs word finally before a pause³¹⁶, prior to morpheme boundaries, and before the continuants {*s*, *ṣ*, *ś*, *h* (*r*) *y*, *v*, *l*}, while OIA *m* never occurs in these environments. This suggests a possible junctional interpretation of *ṁ*, which substitutes for *m* and is effectively representing an allophone of *m* in these situations. This is seen in the analysis of morpheme boundaries in the following examples: Skt. *samrāt* = *samrāt*; *saṁrāddha*- = *sam* + *rāddha*-; *tām sam* = *tām* + *sam*; *tām hanti* = *tām* + *hanti*; *saṁyata*- = *sam* + *yata*-; *samyak* = *samyak*; *amlāna*- = *amlāna*-, and *saṁlāpa*- = *sam* + *lāpa*-. Ongoing developments in OIA led to the analogical use of *ṁ* even within words and morphemes, such as *saṁvṛta*- for example, although such uses are not always consistently applied. The treatment of *anusvāra* as a phonemic nasal consonant is supported by pairs such as *vaṁśa*- (MW. s.v. m. "reed") and *vaśa*- (m. "power, influence"), and *māṁsa* (n. "flesh") and *māsa* (m. "month") for example. And while *anusvāra* has been described as a pure nasal sound, it is also suggested that when appearing before palatal *ś* or velar *h* it represents the corresponding class nasal, as in *ñ* or *ṇ*, and before dental *s* it represents either *m* or *n*.

Although the early recording of *anusvāra* may have originated from the perception of post-vocalic nasality or 'after sound' in certain environments, this nasality appears to have acquired a phonemic consonantal status. This in turn led to an extended orthographic use of the *anusvāra* sign to represent nasal consonants in most pre-consonantal situations, where they are largely allophonic in that they assimilate the place of following consonants. In this context it appears that the *anusvāra* symbol *ṁ* evolved from the marking of vowel nasality to serve as a convenient broad phonemic representation of pre-consonantal nasals.

³¹⁶ Texts normally have *m* before a pause, but this is represented by *anusvāra* in sandhi rules.

Appendix A3 - Sanskrit consonant clusters

A tabulation of OIA consonant clusters is presented in Table A3.1. This has been developed from available listings of clusters in both Vedic (Kobayashi 2004: 185-191 [§125 -§129]) and Classical Sanskrit (Macdonell 1927: 6-8 [§13]), which overlap in most but not all cases, along with a small number of clusters listed in other sources. While these listings are comprehensive, other less common clusters may also occur in the literature. The table includes both two and three consonant clusters, and has been arranged on the basis of both manner and place of articulation for the initial and final consonant in each cluster, along with word initial and medial occurrences. Approximately 277 two consonant medial clusters are listed, of which some 81 also occur word initially, and a further 3 clusters are reported as word initial only (*chm-*, *chn-* & *chl-*), giving a total of 280 clusters.

A further 71 clusters have three or more consonants, and these have been listed separately in Table A3.2. All of these are extensions of the two consonant clusters, mostly by the addition of a final liquid or glide, but also in four cases with an additional nasal or fricative, as in {*kṣ-*, *ñc-*}-*m* and *ṅk-*{*-ṣ*, *-ṣv*}. The initial paired clusters have varying sonority patterns, but in all cases the final consonant is more sonorant than the preceding element.

Table A3.1: Sanskrit Consonant clusters

		Stops																		Nasals					Fricatives				Liquids		Glides									
		[- cor, + ant, -/+ voice]				[+ cor, + ant, -/+ voice]				[+ cor, +/- ant, -/+voice]				[+ cor, - ant, -/+ voice]				[- cor, - ant, -/+ voice]				[+ voice]					[- voice]				[+ voice]		[+ voice]							
Geminates		Labial (L)				Dental-Alveolar (D)				Retroflex (R)				Palatal (P)				Velar (V)				L	D	R	P	V	D	R	P	G	(A)	D	P	L						
		C =	<i>p</i>	<i>p^h</i>	<i>b</i>	<i>b^h</i>	<i>t</i>	<i>t^h</i>	<i>d</i>	<i>d^h</i>	<i>ṭ</i>	<i>ṭ^h</i>	<i>ḍ</i>	<i>ḍ^h</i>	<i>c</i>	<i>c^h</i>	<i>j</i>	<i>j^h</i>	<i>k</i>	<i>k^h</i>	<i>g</i>	<i>g^h</i>	<i>m</i>	<i>n</i>	<i>ṇ</i>	<i>ṅ</i>	<i>ṁ</i>	<i>s</i>	<i>ʃ</i>	<i>ś</i>	<i>h</i>	<i>r</i>	<i>l</i>	<i>y</i>	<i>v</i>					
Stops	Labial	<i>p</i> C-	r																													R	r	r						
		- <i>p</i> C-	r w				R w	pty								V								V m	R w	r			R w	psv		R		R w	r w	R w	(r) w			
		<i>p^h</i> C-																																						
		- <i>p^h</i> C-																																						
		<i>b</i> C-																															R							
		- <i>b</i> C-				w	V w											R w					w										R w	V	(r) w	(r) w				
		<i>b^h</i> C-																															R							
		- <i>b^h</i> C-																															R w	V	R w	R w				
	Dental-Alveolar	<i>t</i> C-																						R									r			R		R	R	
		- <i>t</i> C-	R w					R w	tty	R w											R w	tkr	V		R w	tmy	R w	tny					R s			R w	try		R w	R w
		<i>t^h</i> C-						ttr																									tsny							
		- <i>t^h</i> C-																																			r		R w	R
		<i>d</i> C-																																R			R	R		
		- <i>d</i> C-				R w	R w																		R w		R w							R w	dry		R w	R w	dv	
		<i>d^h</i> C-																							r									R		r	R			
		- <i>d^h</i> C-																							R w		R w							R w	dhry		R w	R w		
	Retroflex	<i>ṭ</i> C-																																					R	
		- <i>ṭ</i> C-	V					r								R w									V			V					V	V	V		R		R w	r
		<i>ṭ^h</i> C-																																						
		- <i>ṭ^h</i> C-																																			w		w	

			Labial				Dental-Alveolar				Retroflex				Palatal				Velar				L	D	R	P	V	D	R	P	G	(A)	D	P	L				
		C =	p	p ^h	b	b ^h	t	t ^h	d	d ^h	ʈ	ʈ ^h	ɖ	ɖ ^h	c	c ^h	j	j ^h	k	k ^h	g	g ^h	m	n	ɳ	ɳ̃	ṇ	s	ʃ	ś	h	r	l	y	v				
Stops (ctd)	Retroflex (ctd)	ɖ C-																																					
		-ɖ C-			r	r				V			V	R w							R w	w	dgy	w	dghr	w						R		R w	r				
		ɖ ^h C-																																					
		-ɖ ^h C-																														V w		(R)w	R				
	Palatal	c C-																																	R				
		-c C-													R w	w	cchr							r w				V w					r		R w				
		c ^h C-																						r	r							R	r	R	r				
		-c ^h C-																														r w		r w	r				
		j C-																						R								R		R	V				
		-j C-																						R w								R w		R w	R w				
		j ^h C-																																					
		-j ^h C-																																					
	Velar	k C-																															R	r	V	(R)			
		-k C-					R w	R							w						r w	V w			R w	R w	kny	R w		r	R w	V		R w	kry	V w	(R)w	R w	kvy
		k ^h C-																																	R				
		-k ^h C-																									V					w		R w	r				
		g C-																						R	R								R	r					
		-g C-			V	R																		r	R	R	r						R w	gry	V	R w	R		
		g ^h C-																							R								r		V				
		-g ^h C-																						r w	R w	ghny							r w		r w	r			

			Labial				Dental-Alveolar				Retroflex				Palatal				Velar				L	D	R	P	V	D	R	P	G	(A)	D	P	L	
			C =	p	p ^h	b	b ^h	t	t ^h	d	d ^h	ṭ	ṭ ^h	ḍ	ḍ ^h	c	c ^h	j	j ^h	k	k ^h	g	g ^h	m	n	ṇ	ñ	ṅ	s	ś	ṣ	h	r	l	y	v
Nasals [+ voice]	L	m C-																													r	r	r			
	L	-m C-	R w		R w	R w			-ṃd-	r													R w	R w	R						R w	r w	R w	(R) w		
	D	n C-						V																									R	(R)		
	D	-n C-	w npr				R w	R	R w	R w ndhr				Nasals also recorded as ṃ in many texts									R w	R w				w				w		R w	R	
	R	-ṇ C-									w	w	R w ṇḍy	V w					ṅkt ṅkty				r w		V w								R w	R w		
	P	-ñ C-											R w ñcm	r w	R w ñjy				ṅkṣ ṅkṣv														V	(r)		
	V	-ṅ C-							r					ñcy					R w ṅky	R w ṅkhy	R w ṅgy	R w ṅghy	V w	V w			w						w			
Fricatives [- voice]	D	ś C-	R	r			R	R										R					R	r							R		R	R		
	D	-ś C-	R w	R w			R w	R w										R w	V w				R w smy	R w sny				R w				R w		R w	R w	
	R	ṣ C-	V	r						R	R							r					R		r								R	R		
	R	-ṣ C-	R w ṣpr	r						R w ṣṭy	R w							R w	V				R w		R w ṣṇy					R		r	V	R w	R w	
	P	ś C-							ṣṭr	ṣṭry	ṣṭv			R								r	R								R	R	R	R		
	P	-ś C-	R											R w ścy									R	R w					w		R w	R w	R w	R w śvy		
	G	h C-																						r								R	r	R	R	
	G	-h C-																					R w	R w	r w							R w	V w	R w	R w	
Liquid	(A)	-r C-	r x	R	r	r x	R x	R x	R x	R x	V		R		R x	V	R x		R	V x	R x	R	r x rmy	R x	R x	r		r	R x	R x	R		V	R x	R x	
	D	-l C-	R w	r	R	V			r									r w		R			r w						r	V w		r w	R w	r w		
Glide	P	-y C-																																R w	r w	
	L	-v C-																						R w	R				w	R w wi	r	R w wi		w		
Key to source: Kobayashi (RigVeda): R = RV > 10 times; r = RV < 10; V = other Veda. Macdonell: w. Other sources: x																											wi: also occurs word initially									
Multiple conjuncts (3 or more) are typically from the Macdonell listing.																																				

Source: compiled from Macdonell (1927: 6-8) and Kobayashi (2004: 185-191)

Table A3.2: Three or more consonant clusters in Sanskrit

# Equal sonority +	<i>jñ-y</i>	fricative-approximant +
stop-stop +	<i>tn-y</i>	<i>śr-y</i>
<i>kt-r kt-ry kt-v kt-y</i>	<i>tm-y</i>	<i>śv-y</i>
<i>kth-y</i>	<i>dhn-y</i>	# Decreasing sonority +
<i>cch-r</i>	stop-fricative +	nasal-stop +
<i>ḍg-y</i>	<i>kṣ-m kṣ-v kṣ-y</i>	<i>ṅk-t ṅk-ty ṅk-y ṅk-ṣ ṅk-ṣv</i>
<i>ḍgh-r</i>	<i>ts-n ts-ny</i>	<i>ṅkh-y</i>
<i>tk-r</i>	<i>ps-v</i>	<i>ṅg-y</i>
<i>tt-y tt-r tt-v</i>	stop-approximant +	<i>ṅgh-r ṅgh-y</i>
<i>tp-r</i>	<i>kr-y</i>	<i>ñc-m ñc-y ñj-y</i>
<i>dg-r</i>	<i>kv-y</i>	<i>ṇḍ-r ṇḍ-ry ṇḍ-y</i>
<i>dgh-r</i>	<i>gr-y</i>	<i>nt-r nt-y</i>
<i>dd-y</i>	<i>tr-y</i>	<i>nd-r</i>
<i>ddh-y</i>	<i>dr-y</i>	<i>ndh-r</i>
<i>dbh-y</i>	<i>dv-y</i>	<i>np-r</i>
<i>pt-y</i>	<i>dhr-y</i>	<i>mp-r</i>
<i>bbh-y</i>	fricative-nasal +	fricative-stop +
# Increasing sonority +	<i>ṣṇ-y</i>	<i>śc-y</i>
stop-nasal +	<i>sn-y</i>	<i>ṣṭ-r ṣṭ-ry ṣṭ-v ṣṭ-y</i>
<i>ghn-y</i>	<i>sm-y</i>	<i>ṣp-r</i>
# Clusters are grouped according to sonority of initial pair		<i>st-y</i>

Source: compiled from Macdonell (1927: 6-8)

While other longer clusters not listed in the above sources appear to be less common, five consonant clusters are found in the words *kārtsnya* and *tārksṇya*, while others such as these quite possibly occur in the literature.

Appendix A4: Sanskrit syllable patterns

The Sanskrit two consonant clusters listed in Table A3.1 have been categorised in terms of the relative sonorities of the consonants (§ 4.4) and their respective manners of articulation, and the number of clusters falling within these categories are listed in Table A4.1. The clusters falling within each grouping are described in more detail below.

Table A4.1: Sanskrit two consonant clusters in terms of manner and relative sonority

Sonority pattern by manner	Medial	Initial
Increasing sonority	Total 137	Total 66 repeat & 3 exclusive
Stop: - fricative	9	3
- nasal	30	7 + 2 excl. (<i>chm-</i> & <i>chn-</i>)
- liquid (<i>r, l</i>)	24	14 + 1 excl. (<i>chl-</i>)
- glide (<i>y, v</i>)	34	17
Fricative: - nasal	9	7
- liquid	7	5
- glide	8	8
Nasal: - liquid	3	2
- glide	9	3
Liquid: - glide	4	--
Equal sonority	Total 63	Total 3 (2 rare)
Stop: - stop (incl. geminates) (typically, voiceless-voiceless or voiced-voiced)	44	2 (<i>pp-</i> & <i>gdh-</i> occur only in RV and appear to be rare)
Fricative: - fricative	3	--
Nasal: - nasal	10	--
Liquid - liquid	2	--
Glide: - glide	4	1 (<i>vy-</i>)
Decreasing sonority	Total 77	Total 12
Fricative: - stop	14	11
Nasal: - stop	22	--
- fricative	1	--
Liquid: - stop	24	--
- fricative	6	--
- nasal	5	--

Glide: - stop	--	--
- fricative	1	--
- nasal	2	--
- liquid	2 (-vl-, -vr-)	1 (vr-)

OIA clusters having **equal sonority** such as stop-stop, fricative-fricative, nasal-nasal, liquid-liquid and glide-glide do not occur word initially in the majority of cases, and in the absence of gemination, it is assumed that these will divide as coda-onset pairs, as in *bhak.ti-* and *jan.man-* for example. Three exceptions are *pp-* and *gdh-*, which are reported in the Ṛgveda and appear to be infrequent, and *vy-* which can also be read as *vi-* due to the vocalic nature of the glide *y*. Geminate consonants also divide as coda and onset across the syllable boundary (Kobayashi 2004: 37), as in *-trṇ.ṇa-* and *vr̥k.ka-* for example, similar to the general cross linguistic pattern (Ladefoged and Maddieson 1996: 92).

OIA clusters having **increasing sonority** occur in most of the manner groupings, and about 50% of the medial clusters also occur word initially and are expected to form syllable onsets in medial positions³¹⁷, while the others are assumed to be assigned as coda-onset pairs. Reasonably high proportions of initial clusters occur in the case of stop + {liquid, glide} (32 of 58), fricative + nasal (7 of 9), fricative + {liquid, glide} (13 of 15) and nasal + {liquid, glide} (5 of 12) clusters, while a smaller proportion occur among the stop + fricative (3 of 9) and stop + nasal (7 of 30) clusters. Word initial stop + C clusters are potentially problematic from an articulatory and perceptual perspective since stops do not possess internal perceptual cues, and rely on formant transitions from and to adjacent vowels in order to be perceived distinctly (Davenport and Hannahs 2010: 68-69). The release bursts and formant transitions are the most dominant cue, and in word initial clusters these will be modulated in some way by the following consonant, and the ease with which the stop can be identified in this transition appears to depend upon the manner of the second consonant. Based on the relative proportions seen in word initial positions, the stop + semi-vowel clusters appear to be the most easily realised, while the stop + fricative or nasal clusters appear to be more problematic. Since the semi-vowels are more open and vowel like in their nature, and similarly possess internal formant cues (Davenport and Hannahs 2010: 66), it appears that the formant transition and release

³¹⁷ As discussed in § 4.5, the assessment of syllable weights in metrical texts dictates placement of the first consonant in the cluster in the prior syllable, which may lead to a doubling of this consonant to satisfy sonority sequences in cases such as these.

bursts of the stops are able to find expression in the transition to the following semivowel. At the same time there appears to be an opportunity for epenthetic vowels to arise in this articulatory transition, and this tendency has been recorded in some of the Gāndhārī reflexes (§ 10.2.1). The word initial stop + fricative and stop + nasal clusters have some peculiar features of their own, which have a bearing on how they develop in Gāndhārī, and these are addressed in more detail below:

- Stop + fricative (3 initial cases out of 9 medial): of the initial clusters, *ps*- and *ts*- appear to have limited application, as seen in the small numbers of these listed in Monier Williams (1899), while use of the *kṣ*- cluster is reasonably common, as for example in *kṣaṇa*- (MW s.v. m “moment”) and *kṣapā*- (f “night”). Similar to the semi-vowels, fricatives possess strong internal cues in the form of frication, while the common lexical occurrence of the *kṣ*- cluster in OIA suggests that in the absence of a preceding vowel the velar stop must rely on some form of perceptible transitional cues in its release through the fricative. While the retroflexion of PIIr *ś* > OIA *ṣ* / *k* _ is well established in the RUKI rule (Appendix A1), the release of word initial velar *k* through the retroflex fricative *ṣ* appears to reflect an additional upward movement of the tongue tip following release of the stop. An alternative realisation as palatal *ś*, involving the positioning of the tongue blade near the palate, may involve less articulatory effort and is possibly more compatible with the release of velar *k*. Although **kś*- does not occur initially in contrastive situations, it does occur word medially in OIA. It has also been suggested that the OIA /TS/ clusters may have acquired the status of single phonemes in the form of affricates (Kobayashi 2004: 59-60). This is more likely to occur, however, when frication can arise more naturally during release of the stop in a homorganic cluster, as in */ts-/* [t͡s] for example (Ladefoged and Maddieson 1996: 90-91), and **kś* [k͡ɕ] is not inconsistent with this interpretation. The weakening of Skt. *kṣ*- > *kh*- in Pali also points towards a developing process of debuccalisation (Fallon 2002: 123-127). Medial */TS-/* clusters are also quite common in OIA (Kobayashi 2004: 59), seen in *akṣara* (MW s.v. n “syllable”), *mat-sara* (adj. “joyous”) and *ap-saras*- (f “female divinities”) for example, where the stop can gain some definition through the transitional formants from the preceding vowel. It is unclear to what extent the *ts* and *ps* clusters may occur as syllable onsets due to their limited word initial status.

- Stop + nasal (7 of 30): nasals are more sonorous than stops due to the open voiced nasal airflow. However, like stops they are non-continuants, and the pronunciation of word initial stop-nasal clusters requires a mid-cluster adjustment between the manners of articulation in which the stop may only be partially released, although in medial positions it can be identified by the weaker V₁C₁ formant transitions (Jun 2004; Wright 2004). This contrasts with homorganic nasal-stop clusters where the nasal sound can be held as such prior to a single oral release. These difficulties of articulatory co-ordination may inhibit the development of stop-nasal clusters as medial onsets, as for example in English ‘atmosphere’ [æt.mə.sfiə]. Hence although listed as word initial, the Vedic clusters *gn-*, *gm-*, *ghn-*, (*chm-* & *chn-*), *jm-*, *tm-* and *dhm-* all appear to be quite rare (Kobayashi 2004: 188-191). And while *jñ-* is common in the many derivations of the verbal root *jñā*, as in *jñā.na-* and *abhi.jñā-* for example, other homorganic clusters such as *[*tn-*], *[*tl-*] and *[*pm-*]³¹⁸ are disallowed as onsets in the *[place][place] language specific condition (Cho 1999: 62-64, 159), suggesting *jñ-* is an exception. Pronunciation of the *jñ-* cluster as a unitary segment appears to be quite difficult without some form of intervening release, which may be realised as a short epenthetic vowel or frication for example, and pronunciations in Modern IA languages such as ‘gy’ in Hindi, or ‘gn’ in Gujarati appear to resolve these articulatory difficulties in different ways (Kobayashi 2004: 52-53). Stop-nasal clusters also develop as unitary ‘pre-stopped’ nasal segments in many languages, with little distinction from a single nasal (Ladefoged and Maddieson 1996: 128-129). It is therefore possible that in word initial *jñā-* the *j-* may simply serve as an orthographic place signal for the following nasal, and the cluster pronounced as a unitary [j̥n] segment. This development may also apply in word medial positions, or alternately the cluster may develop as coda-onset where the initial stop can be identified more clearly through its formant transitions from the preceding vowel. Other medial clusters commonly form across morpheme boundaries, seen in *tṛp-no-ti* for the verb √*tṛp* (MW s.v. “be satisfied”) and *ud-naya-ti* (realised as *unnayati*) for the prefixed verb √*nī* (“lead”) suggesting a similar syllable division.

OIA clusters with **decreasing sonority** typically only occur word medially, except for fricative + stop clusters, in which 11 of the 14 medial clusters also occur initially in

³¹⁸ Although written as *[*tm-*] on page 62 in (31b) it appears from the context of the discussion on page 64 that this is intended to represent the homorganic cluster *[*pm-*].

violation of the sonority sequencing principle. These types of clusters are expected to divide as coda and onset pairs, and it is also suggested that medial /s/ + stop clusters remain hetero-syllabic (Cho 1999: 62-63). Inconsistencies of this type have necessitated the inclusion of ‘exceptions’ which are problematic for attempts to provide a unified characterisation of syllable structure using sonority sequencing (Wright 2004; Henke et.al. 2012). One approach is to classify word initial sibilants which occur prior to voiceless stops as extra-syllabic sibilants, represented as <S>T_[0voiced]-, as in *skand^ha* = /<s>kan.d^ha-/ and *ścandra*- = /<ś>can.dra- for example (Kobayashi 2004: 41-43). Gordon (2016: 103-104) observes that sibilants are largely free of the sonority sequencing constraints of other sounds, and to accommodate such patterns an alternative assessment of sonority or strength sequencing is proposed, based on the robustness of the perceptual cues of consonants. In this instance it is suggested that sibilant fricatives have strong internal cues which protect their identity in pre-consonantal position, including word initially, where they do not need to depend upon formant transitions from a preceding vowel to be distinctly perceived (Wright 2004: 50; Gordon 2016: 116-117). This may also be thought of in terms of articulatory ease, where the closure to an initial stop is able to transition through the partial closure of a fricative with little additional articulatory effort, similar to the fricative release of stops mentioned above. The majority of the other clusters with decreasing sonority belong to the nasal + stop (22) and liquid + stop groupings (24), which involve the liquid consonants in –rT- (17) and –lT- (7) clusters. Nasals have strong cues to voice and manner, while their place cues are weak (Wright 2004: 50), and are able to survive as codas by generally assimilating to the place of following stops.

The OIA three consonant clusters typically follow the same general patterns of syllabification. Since the final consonants are more sonorant in all cases, the clusters listed in Table A3.2 will typically pattern as a coda plus two consonant onset.

Appendix B: Perceptual cues of consonants

The main auditory cum perceptual cues of consonants based on the manner of articulation are summarised in Table B1. The manner cues are those which help listeners identify the manner of the consonant, while the place cues are those which help distinguish the place of articulation of particular consonants in that manner group. Internal cues are any acoustic signals emitted during the period of articulation of the consonant, while transitional cues are those that occur in the transition time between articulation of the consonant and any preceding or following vowel or consonant.

Table B1: The main perceptual cues of consonants

	Manner cues &	Place cues	
	Type of attenuation (1)	Internal	Transitional
Stops (T)	Attenuation is abrupt in all frequencies except F0. Period of silence (voiced may have some residual). Formant transitions are the most abrupt.	None	Release bursts most dominant. Formant transitions from / to adjacent vowels, especially F2 and to lesser degree F3, also play a role.
Nasals (N)	Nasal murmur. Less severe drop in amplitude. Look ahead cues such as nasalisation of lead in V.	Nasal resonance formants (anti-formants), but not reliable.	F2 formant transitions during oral release appears to be more powerful and reliable in terms of place.
Fricatives (S)	Complete attenuation, but accompanied by fricative noise. Higher intensity frication is more reliable.	Fricative noise spectrum is the primary cue, and enables perception of different places of sibilant fricatives.	F2 formant transitions from / to adjacent vowels, also play a role in distinguishing place.
<i>l</i>	The approximants (R), comprising liquids and glides, maintain formant structure. Glides (y,v) also have more gradual formant transitions.	Liquids and glides are distinguished from each other on the basis of their internal formant values at the peak of constriction.	Also dependent on the presence of a neighbouring vowel to carry transitional information.
<i>v</i>			
<i>y</i>			
<i>r</i>			
(1): All oral constrictions result in attenuation of the signal, especially in higher frequencies, and the relative degree of attenuation is a strong cue to the manner consonants.			

Source: Compiled from Wright (2004: 36-40)

Formants are resonances of the fundamental frequencies of vibration of the vocal cords, which develop in the vocal tract according to the placement of the articulators. Vowel sounds typically display a number of formants above the fundamental frequency, which are labelled as F1, F2, F3 and so forth. These provide a unique identity to each of the vowel sounds, while more open sonorant consonants such as approximants also have formant patterns. Vowel formant frequencies display transitional patterns as they merge into or away from adjacent obstruent consonants, which correlate with and help identify the place of articulation of the consonant (Davenport and Hannahs 2010: 61-71).

The main points to be drawn from Table B1 are summarised as follows:

- Stops (T) have no internal formants and are reliant on a period of silence followed by release bursts for their manner identity, and formant transitions to adjacent vowels for place identity.
- The manner identity of nasals (N) is reliant on the nasal murmur as well as anticipatory nasalisation. The internal formants are confused by anti-formants and a poor distribution of internal energy, making their place clues reliant on the following formant transitions, although the anti-formants may vary depending upon place of articulation.
- Sibilant fricatives (S) display internal frication and noise which provide reliable internal cues to their manner identity. Variations in the noise spectrum provide reliable indicators of place. Non-sibilant fricatives are reliant on the formant transitions to adjacent vowels, although glottal fricatives can tend to take on features of the adjacent vowels.
- Approximants (R) have strong internal formants like vowels, which provide reliable cues to the manner and place of articulation, supported by more gradual formant transitions to adjacent vowels.

Appendix C: Comparative reflexes of Sanskrit consonants in Aśokan inscriptions

Table C1: Comparative reflexes: Aśokan intervocalic single consonants

Sanskrit	Aśokan Inscriptions 3 BCE			G Buddhist Manuscripts § 8.2	Literary Prakrits - Śaurasenī
	North West	East	West		
<i>k</i>	<i>k, y</i>	<i>k, g, y</i>	<i>k</i>	<i>g, Ø</i>	<i>y / Ø</i>
<i>kh</i>		<i>h</i>	<i>kh</i>	<i>h</i>	<i>> gh > h</i>
<i>g</i>	<i>g, k, y</i>	<i>g, k, y</i>	<i>g, y</i>	<i>g, g, Ø</i>	<i>y / Ø</i>
<i>gh</i>		<i>h</i>	<i>h</i>	<i>g, h, y</i>	<i>gh > h</i>
<i>ṇ</i>				<i>ṃ</i>	<i>ṃ</i>
<i>c</i>	<i>c</i>	<i>c, j</i>	<i>c</i>	<i>y / Ø</i>	<i>y / Ø</i>
<i>ch</i>					<i>(c)ch</i>
<i>j</i>	<i>j, y</i>	<i>j, c</i>	<i>j, c</i>	<i>y / Ø</i>	<i>y / Ø</i>
<i>jh</i>					<i>h</i>
<i>ñ</i>				<i>ṃ</i>	<i>ṃ</i>
<i>ṭ</i>	<i>ṭ</i>	<i>ṭ</i>	<i>ṭ</i>	<i>ḍ</i>	<i>ḍ</i>
<i>ṭh</i>		<i>ṭ</i>		<i>ḍh</i>	<i>ḍh</i>
<i>ḍ</i>		<i>ḍ, ḷ</i>		<i>ḍ</i>	<i>ḍ</i>
<i>ḍh</i>			<i>ḷ</i>	<i>ḍh</i>	<i>ḍh</i>
<i>ṇ</i>	<i>ṇ</i>	<i>n</i>	<i>ṇ</i>	<i>ṇ</i>	<i>ṇ</i>
<i>t</i>	<i>t, d</i>	<i>t</i>	<i>t</i>	<i>ḍ</i>	<i>> d > y / Ø</i>
<i>th</i>		<i>dh</i>		<i>ṣ</i>	<i>h</i>
<i>d</i>	<i>d</i>	<i>d, t</i>	<i>d, Ø</i>	<i>ḍ</i>	<i>> y / Ø</i>
<i>dh</i>	<i>dh</i>	<i>dh, h</i>	<i>dh, (d)</i>	<i>ṣ</i>	<i>h</i>
<i>n</i>					<i>ṇ</i>
<i>p</i>	<i>p, v</i>	<i>p, v</i>	<i>p</i>	<i>v / Ø</i>	<i>> b > v / Ø</i>
<i>ph</i>					<i>h</i>
<i>b</i>				<i>v / Ø</i>	<i>> v / Ø</i>
<i>bh</i>	<i>bh, h</i>	<i>bh, h</i>	<i>bh, h</i>	<i>bh / h</i>	<i>h</i>
<i>m</i>	<i>m</i>	<i>m, (ph)</i>	<i>m</i>	<i>m</i>	<i>m</i>
<i>y</i>	<i>y, j, Ø</i>	<i>y, j</i>	<i>y</i>	<i>y</i>	<i>Ø</i>
<i>r</i>	<i>r</i>	<i>l</i>	<i>r</i>	<i>r</i>	<i>r</i>
<i>l</i>	<i>l, (ḍ)</i>	<i>l, (ḍ)</i>	<i>l, (ḍ)</i>	<i>l</i>	<i>l</i>
<i>v</i>				<i>v</i>	<i>v</i>
<i>ś</i>	<i>ś</i>	<i>s</i>	<i>s</i>	<i>ś</i>	<i>s</i>
<i>ṣ</i>	<i>ṣ</i>	<i>s</i>	<i>s</i>	<i>ṣ</i>	<i>s</i>
<i>s</i>	<i>s</i>	<i>s</i>	<i>s</i>	<i>s</i>	<i>s</i>

Source: compiled from Hultzsch (1925: lvii-lviii, lxxxv-lxxxvi, c-ci), Mehendale (1948:

210-216 [Chapter IV, Synoptic Tables]), Misra (1967: 162-172)

Table C2: Comparative reflexes: Aśokan medial consonant clusters

	Aśokan Inscriptions 3 BCE					
Sanskrit	North West Shāhbāzgarhī		East Dhauḷi and Jauguḍa		West Girnār	
Hultzsch 1925:	lxxxvii-xc		cii -civ		lviii-lxii	
Mehendale 1948 (xxiv-xxv, 21-27, 217-237 [Chapter IV, Synoptic Tables])						
Stop and nasal combinations of occlusives						
	Hultzsch	M'dale	Hultzsch	M'dale	Hultzsch	M'dale
T ₁ T ₂	T ₂ T ₂ (typical)		T ₂ T ₂		T ₂ T ₂	
kt	tt		tt		tt	
tp	pp		pp		tt	
dg			gg			
pt	tt		tt			
bdh	ddh				ddh	
TN	TT typ only - varies		TT typ only		TT typ only	
gn		gg	gg / gin	gg	gg	gg
jñ		ñ / ṇ /n / (jin)	jīn / mñ / ṇ	ñ / mñ	ñ / mñ	ṇ / mn / mñ / (jin)
tm	tv / tt	tt	tt	tp	tp	tt
pn		puṇ	pun	puṇ	pun	pun
NT	ṁT		ṁT		ṁT	
nḍ		mḍ				mḍ
ñc	mc	mc		mc		mc / mn
ñj	mj	j / (ñ)		mj		mj
ndh	mndh / dh					
NN						
Clusters having rising sonority (falling strength)						
TS	TT / SS		TT / (S)		TT / (S)	
kś			kkh		cch	
kṣ		kkh		kkh	kh / cch	kh
ts		s	s	ch	cch	s
ts				ch		
TI			TI / TT		TI	
kl			kil			kil
khl			kkh			

Tv	Tv / T		Tv		Tv / Tp	
<i>tv</i>	<i>tv / t</i>	<i>t</i>	<i>tv</i>	<i>tp</i>	<i>tp</i>	<i>t</i>
<i>dv</i>			<i>v / duv</i>		<i>dv / db</i>	<i>duv</i>
Ty	TT typ		TT typ		TT typ	
<i>ky</i>	<i>kk</i>		<i>kiy</i>		<i>kk</i>	
<i>khy</i>	<i>kkh</i>	<i>kh</i>	<i>khy / khiy</i>	<i>ch</i>		<i>kkh</i>
<i>jy</i>	<i>jj</i>	<i>j</i>	<i>jj</i>			<i>jj</i>
<i>ty</i>	<i>cc</i>	<i>tiy</i>	<i>tiy</i>	<i>c</i>	<i>cch</i>	<i>tiy / ty</i>
<i>dy</i>	<i>yy</i>	<i>diy</i>	<i>y / j</i>	<i>y / j</i>	<i>dd / jj</i>	<i>diy / jj</i>
<i>dhy</i>		<i>dhiy</i>	<i>jh</i>	<i>jh</i>	<i>jh</i>	<i>jh / dhy</i>
<i>py</i>					<i>pp</i>	
<i>bhy</i>	<i>bbh</i>		<i>bhiy</i>		<i>bbh</i>	
Tr	Tr		TT		TT	
<i>kr</i>	<i>kr</i>	<i>kr</i>	<i>kk</i>	<i>kr</i>	<i>kr / k</i>	<i>kk</i>
<i>gr</i>	<i>gr / ghr</i>	<i>gr</i>	<i>gg</i>	<i>gg</i>	<i>gg</i>	<i>gg</i>
<i>tr</i>	<i>tr</i>	<i>tr</i>	<i>tr / tt</i>	<i>tr</i>	<i>tr / tt</i>	<i>tt</i>
<i>dr</i>	<i>dr</i>	<i>dr / d</i>	<i>dd</i>	<i>dd</i>	<i>dd</i>	<i>dd</i>
<i>dhr</i>	<i>dhr</i>		<i>ddh</i>		<i>dhr</i>	<i>ddh</i>
<i>pr</i>	<i>pr</i>		<i>p</i>		<i>pr / p</i>	
<i>br</i>			<i>bb</i>			
NS						
Nv						
Ny	NN / (Ny)		NN / (Ny)		NN / (Ny)	
<i>ny</i>	<i>ññ / ññ</i>	<i>ññ / ñ</i>	<i>my / ñn</i>	<i>ññ / ñ</i>	<i>ññ / ñ</i>	<i>ñn / n / (ññ)</i>
<i>ny</i>	<i>ññ / niy</i>	<i>niy / ñ</i>	<i>ny / ñn</i>	<i>ñn / ññ</i>		<i>niy / ñn</i>
<i>my</i>	<i>my / mm</i>		<i>my</i>		<i>my</i>	
Nr						
<i>mr</i>	<i>mb</i>				<i>mb</i>	
Sl						
<i>sl</i>						<i>s</i>
Sv			Sv		Sv	
<i>sv</i>			<i>sv</i>		<i>sv</i>	<i>sv</i>
<i>sv</i>		<i>sp-?</i>	<i>sv</i>	<i>sv</i>	<i>sv / s</i>	<i>s, suv</i>
Sy	SS / (Sy)		Sy / (SS)		SS / (Sy)	
<i>sy</i>	<i>siy</i>		<i>s</i>		<i>siy / s</i>	
<i>sy</i>	<i>śś</i>				<i>s</i>	
<i>sy</i>	<i>siy / ss</i>	<i>siy / siy</i>	<i>sy / siy</i>	<i>siy</i>	<i>s</i>	<i>siy / s / sy</i>

Sr	Sr / SS		SS		Sr / SS	
<i>śr</i>	<i>śr / śś</i>		<i>s</i>		<i>sr / s</i>	
<i>sr</i>	<i>sr / ss</i>	<i>sr</i>	<i>s</i>	<i>s</i>	<i>sr / s</i>	<i>s</i>
lv						
ly	ll		y / l		l / y	
<i>ly</i>	<i>ll</i>	<i>l</i>	<i>y</i>	<i>l</i>	<i>ll</i>	<i>y</i>
vy	vv / vy		vy		vy / vv	
<i>vy</i>	<i>vv</i>	<i>viy / v</i>	<i>vy / viy / y</i>	<i>vy</i>	<i>vy / v</i>	<i>viy /vy / y</i>
vr	vr		vv		vv	
<i>vr</i>	<i>vr</i>	<i>vr</i>	<i>v</i>	<i>v</i>	<i>v</i>	
Clusters having decreasing sonority (increasing strength)						
ST	TT(h)		TT(h)		TT(h)	
<i>śc</i>	<i>cc</i>	<i>c</i>		<i>ch</i>		<i>ch</i>
<i>śch</i>			<i>chk ?</i>		<i>cch</i>	
<i>śp</i>			<i>pph</i>			<i>ph / p</i>
<i>śk</i>	<i>kk</i>		<i>kk</i>		<i>kk</i>	
<i>śt</i>	<i>t̥th</i>	<i>t̥h</i>				<i>t̥h</i>
<i>śth</i>	<i>t̥th</i>	<i>t̥h / th</i>	<i>t̥h / th</i>	<i>st̥</i>	<i>st̥</i>	<i>t̥h</i>
<i>sk</i>	<i>kk</i>	<i>k</i>		<i>k</i>	<i>kh</i>	<i>k</i>
<i>st</i>	<i>st̥</i>	<i>st̥</i>	<i>tth / t̥th</i>	<i>st / (t̥h)</i>	<i>st̥</i>	<i>th</i>
<i>sth</i>	<i>tth</i>	<i>th / t̥h</i>	<i>tth</i>	<i>st̥</i>	<i>th / st</i>	<i>t̥h</i>
<i>hkh</i>			<i>kkh</i>			
SN	SS / Sp		Mixed		mh	
<i>śm</i>			<i>pph</i>			
<i>sm</i>	<i>sp / ss</i>	<i>sp</i>	<i>sm / ph</i>	<i>mh</i>	<i>mh</i>	<i>s</i>
<i>hm</i>	<i>mm</i>	<i>m</i>	<i>mbh</i>	<i>hm</i>	<i>mh</i>	<i>bh ?</i>
IT	lp		pp		pp	
<i>lp</i>	<i>lp</i>		<i>pp</i>		<i>pp</i>	
IN						
IS						
vN						
rT	rT		TT		TT	
<i>rk</i>			<i>kkh</i>			
<i>rg</i>	<i>gr</i>		<i>gg</i>		<i>gg</i>	
<i>rgh</i>					<i>ggh</i>	
<i>rt</i>	<i>tr / tt</i>	<i>t̥ / t̥r / t</i>	<i>tt / t̥t̥</i>	<i>t̥ / t</i>	<i>t / t̥</i>	<i>t̥ / t</i>

<i>rth</i>	<i>rth / thr / tth</i>	<i>ṭh / th / ṭhr / thr</i>	<i>tth / ṭṭh</i>	<i>th</i>	<i>tth</i>	<i>ṭṭh / tth</i>
<i>rd</i>				<i>d</i>	<i>dd</i>	<i>d</i>
<i>rdh</i>	<i>ḍḍh</i>	<i>ḍh / dh / dhr</i>	<i>ddh</i>	<i>ḍh / dh</i>	<i>ddh / ḍḍh</i>	<i>ḍh</i>
<i>rbh</i>	<i>rbh</i>		<i>bbh</i>		<i>bbh</i>	
rN	NN		NN		NN	
<i>rṇ</i>	<i>mṇ</i>					<i>mṇ</i>
<i>rn</i>					<i>mn</i>	
<i>rm</i>	<i>rm</i>		<i>mm</i>		<i>mm</i>	
rS	rS / SS		rS / SS		rS / SS	
<i>rs</i>			<i>s / rs</i>			
<i>rś</i>	<i>rś / śś</i>	<i>ś / rś</i>		<i>s / rs</i>	<i>s / rs</i>	<i>s / (rs)</i>
<i>rṣ</i>	<i>rṣ / ṣṣ</i>	<i>ṣ / rs</i>	<i>s</i>	<i>s / rs</i>	<i>s / rs</i>	<i>s</i>
<i>rh</i>	<i>rah</i>	<i>rah / r</i>		<i>rah</i>	<i>rah</i>	
rl						
rv	rv / vv		rv / vv		rv / vv	
<i>rv</i>	<i>ruv / vv</i>	<i>vr (rv) / v</i>	<i>v / luv</i>	<i>rv</i>	<i>rv / v</i>	<i>v, luv</i>
ry	ry / yy		yy		ry / yy	
<i>ry</i>	<i>riy / yy</i>	<i>riy / y</i>	<i>y</i>	<i>y</i>	<i>y / riy</i>	<i>liy / y</i>

Source: compiled from Hultzsche (1925: lvii-lxii, lxxxvii-xc, cii-civ), and Mehendale (1948: xxiv-xxv, 21-27, 217-237 [Chapter IV, Synoptic Tables]).

References

- Allen, W.S. 1953. *Phonetics in Ancient India*. London Oriental Series – Volume 1. London: Oxford University Press.
- Allon, Mark. 2001. *Three Gāndhārī Ekottarikāgama-Type Sūtras. British Library Kharoṣṭhī Fragments 12 and 14. Gandhāran Buddhist Texts 2*. Seattle and London: University of Washington Press.
- . 2007. “A Gāndhārī Version of the Simile of the Turtle and the Hole in the Yoke”. *The Journal of the Pali Text Society*, Vol xxix, pp. 229-62.
- . 2014. “The Senior Kharoṣṭhī Manuscripts.” In Paul Harrison and Jens-Uwe Hartman, eds. *From Birch Bark to Digital Data: Recent Advances in Buddhist Manuscript Research. Papers Presented at the Conference Indic Buddhist Manuscripts: The State of the Field, Stanford, June 15-19 2009*. pp. 19-34. Wien: Österreichische Akademie der Wissenschaften.
- . 2019. “A Unique Gāndhārī Monastic Ledger Recording Gifts by Vima Kadphises (Studies in Gāndhārī Manuscripts 2).” *Journal of the International Association of Buddhist Studies* 42: 1–46.
- . 2020. “A Gāndhārī Saṃyukta-āgama Version of the Discourse on Not-self (P *Anattalakkhaṇa-sutta*, Skt. **Anātmalakṣaṇa-sūtra*).” In Dhammadinnā, ed. *Research on the Saṃyukta-āgama*. Dharma Drum Institute of Liberal Arts Research Series, 201-258. Taipei: Dharma Drum Publishing Co.
- . 2021. *The Composition and Transmission of Early Buddhist Texts with Specific Reference to Sūtras*. Hamburg Buddhist Studies, 17. Bochum: Projekt Verlag.
- Allon, Mark and Ian McCrabb. 2019. *Gandhāran Buddhist Texts*. University of Sydney. <https://gandhari-texts.sydney.edu.au/>
- Allon, Mark and Richard Salomon. 2000. “Kharoṣṭhī Fragments of a Gāndhārī Version of the Mahāparinirvāṇasūtra.” In: Jens Braarvig, ed., *Buddhist Manuscripts, Volume I*. Manuscripts in the Schøyen Collection, I, pp. 243–73. Oslo: Hermes Academic Publishing.
- Allon, Mark, Richard Salomon, Geraldine Jacobsen and Ugo Zoppi. 2006. “Radiocarbon Dating of Kharoṣṭhī Fragments from the Schøyen and Senior Manuscript Collections.” In: Jens Braarvig, ed., *Buddhist Manuscripts, Volume III*. Manuscripts in the Schøyen Collection, III, pp. 279–314. Oslo: Hermes Publishing.

- Bashir, Elena. 2003. "Dardic." In Cardona, George, and Danesh Jain, eds. *The Indo-Aryan Languages*. pp. 818-894. London and New York: Routledge.
- Baums, Stefan. 2009. *A Gāndhārī Commentary on Early Buddhist Verses: British Library Kharoṣṭhī fragments 7, 9, 13 and 18*. PhD Dissertation. Department of Asian languages and Literature, University of Washington.
- . 2012. "Catalog and Revised Texts and Translations of Gandharan Reliquary Inscriptions." In: David Jongeward, Elizabeth Errington, Richard Salomon and Stefan Baums, *Gandharan Buddhist Reliquaries*. Gandharan Studies, Volume 1, pp. 200–51. Seattle: Early Buddhist Manuscripts Project.
- . 2019. *Outline of Gāndhārī Grammar*. Distributed under the Creative Commons BY-ND license. https://stefanbaums.com/baums_grammar_outline.pdf
- Baums, Stefan and Andrew Glass. 2002a. *Gāndhārī Language and Literature*. Accessed September 27, 2022: <https://gandhari.org>
- . 2002b. *Catalog of Gāndhārī Texts*. Accessed September 27, 2022: <https://gandhari.org/catalog>
- . 2002c. *A Dictionary of Gāndhārī*. Accessed September 27, 2022: <https://gandhari.org/dictionary>
- Baums, Stefan, Andrew Glass and Kazunobu Matsuda. 2016. "Fragments of a Gāndhārī Version of the Bhadrakalpikasūtra." In Jens Braarvig, ed. *Manuscripts in the Schøyen collection, Buddhist Manuscripts, Volume IV*. pp. 184-266, 439-455. Oslo: Hermes Publishing.
- Baums, Stefan, Jens Braarvig, Timothy Lenz, Fredrik Liland, Kazunobu Matsuda and Richard Salomon. 2016. "The Bodhisattvapīṭakasūtra in Gāndhārī." In Jens Braarvig, ed. *Manuscripts in the Schøyen collection, Buddhist Manuscripts, Volume IV*. pp. 267-282, 439-455. Oslo: Hermes Publishing.
- Bloch, Jules. 1950. 1950. *Les inscriptions d'Asoka : traduites et commentées*. Collection Émile Senart, huitième volume. Paris: Société d'édition, Les belles lettres.
- Brough, John. 1962. *The Gāndhārī Dharmapada*. London: Oxford University Press. Reprint Delhi: Motilal Banarsidass, 2001.
- Bubenik, Vit. 2003. "Prākṛits and Apabhraṃśa." In Jain, Danesh, and George Cardona, eds. *The Indo-Aryan Languages*. pp. 204-249. London and New York: Routledge.

- Bühler, G. 1889. “Die Shâhbâzgarhi Version der Felsenedicte Aśoka’s.” *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 43: 128–76.
- . 1892. “Aśoka’s Twelfth Rock-Edict according to the Shâhbâzgarhî Version.” *Epigraphia Indica* 1: 16–20.
- . 1894. “Aśoka’s Rock Edicts according to the Girnar, Shâhbâzgarhî, Kâlsî and Mansehra Versions.” *Epigraphia Indica* 2: 447–72.
- Burrow, T. 1937. *The Language of the Kharoṣṭhī Documents from Chinese Turkestan*. (1st paperback ed. 2011). Cambridge: Cambridge University Press.
- . 1940. *A Translation of the Kharoṣṭhī Documents from Chinese Turkestan*. James G. Forlong Fund, vol. XX. London: The Royal Asiatic Society.
- . 1972. *The Sanskrit Language* (3rd ed.). UK, Faber & Faber. Reprint (1st Indian ed.) Delhi: Motilal Banarsidass, 2001.
- . 1973. “The Proto-Indoaryans”. *The Journal of the Royal Asiatic Society of Great Britain and Ireland*, 1973, No. 2 (1973) pp. 123-140
- Butcher, Michael. 2020. *The Emotionology of Anger in Early Buddhist Literature: Through the Lens of a Gāndhārī Verse Text*. PhD Dissertation. University of Washington.
- Campbell, Lyle. 2013. *Historical Linguistics. An Introduction* (3rd ed.). Edinburgh: Edinburgh University Press.
- Cardona, George. 2003. “Sanskrit.” In Cardona, George, and Danesh Jain, eds. *The Indo-Aryan Languages*. pp. 104-160. London and New York: Routledge.
- Cardona, George. 2013. “Developments of nasals in early Indo-Aryan: anunāsika and anusvāra.” *Tokyo University Linguistic Papers (TULIP)* 33 (2013) 3-81.
- Clackson, James. 2007. *Indo-European Linguistics – an Introduction*. Cambridge. Cambridge University Press.
- Clements, G. N. 1990. “The role of the sonority cycle in core syllabification.” In John Kingston and Mary E. Beckman, eds. *Papers in Laboratory Phonology I: Between the grammar and physics of speech*. pp 283-333, Cambridge, Cambridge University Press
- Cho, Young-mee Yu. 1999. *Parameters of Consonantal Assimilation*. Lincom Studies in Theoretical Linguistics; 15. Muenchen: Lincom Europa.
- Coulson, Michael. 1976. *Teach Yourself Sanskrit*. London: Hodder Education

- Cousins, Lance. 1996. "The Dating of the Historical Buddha: A Review Article".
JRAS, Third Series, Vol. 6, No. 1, pp. 57-63.
- Cox, Collett. (forthcoming). *A Gāndhārī Abhidharma Text: British Library Kharoṣṭhī Fragment 28*. Gandhāran Buddhist Texts XX. Seattle: University of Washington Press.
- Cunningham, Alexander. 1877. *Inscriptions of Aśoka*. Corpus Inscriptionum Indicarum Vol. 1. Calcutta: Office of the Superintendent of Government Printing.
- Davenport, Mike and S. J Hannahs. 2010. *Introducing Phonetics and Phonology* (3rd ed.). London and New York: Routledge.
- Denham, Kristin and Anne Lobeck. 2013. *Linguistics for Everyone. An Introduction*. (2nd ed.). Boston, Wadsworth, Cengage Learning.
- Deshpande, Madhav M. 1983. "Pāṇini as a frontier grammarian." *Chicago Linguistic Society, papers from the 19th regional meeting*, 1983. CLS 19, pp 110-116.
- Edgerton, Franklin. 1936. "The Prakrit Underlying Buddhistic Hybrid Sanskrit." *Bulletin of the School of Oriental Studies* 8: 501–516.
- . 1953. *Buddhist Hybrid Sanskrit Grammar and Dictionary, Volume I: Grammar*. William Dwight Whitney Linguistic Series. New Haven: Yale University Press.
- Esposito, Christina M., Sameer ud Dowla Khan, and Alex Hurst. 2005. "Breathy Nasals and /Nh/ Clusters in Bengali, Hindi, and Marathi." *UCLA Working Papers in Phonetics, No 104*. pp 82-106.
- Falk, Harry. 2011. "The Split Collections of Kharoṣṭhī Texts." *Annual Report of the International Research Institute for Advanced Buddhology (ARIRIAB) at Soka University for the Academic Year 2010*. Vol. XIV (2011), pp 13-23. Tokyo: The International Research Institute for Advanced Buddhology.
- . 2015. "A new Gāndhārī Dharmapada (Texts from the Split Collection 3)." *Annual Report of the International Research Institute for Advanced Buddhology (ARIRIAB) at Soka University for the Academic Year 2014*. Vol. XVIII (March 2015), pp 23-62. Tokyo: The International Research Institute for Advanced Buddhology.
- Falk, Harry and Seishi Karashima. 2003. "Five New Kharoṣṭhī Donation Records from Gandhāra." *Silk Road Art and Archaeology* 9: 71–86.
- . 2008. "Another Reliquary Vase from Wardak and Consecrating Fire Rites in Gandhāra." In: Claudine Bautze-Picron, ed., *Religion and Art: New Issues in*

- Indian Iconography and Iconology*, pp. 63–80. London: The British Association for South Asian Studies.
- . 2012. “A first-century *Prajñāpāramitā* manuscript from Gandharā – *parivarta* 1 (Texts from the Split Collection 1).” *Annual Report of the International Research Institute for Advanced Buddhology (ARIRIAB) at Soka University for the Academic Year 2011*. Vol. XV (2012), pp 19-62. Tokyo: The International Research Institute for Advanced Buddhology
- . 2013. “A first-century *Prajñāpāramitā* manuscript from Gandharā – *parivarta* 5 (Texts from the Split Collection 2).” *Annual Report of the International Research Institute for Advanced Buddhology (ARIRIAB) at Soka University for the Academic Year 2012*. Vol. XVI (2013), pp 97-171. Tokyo: The International Research Institute for Advanced Buddhology.
- Falk, Harry and Elisabeth Steinbrückner. 2020. “A Metrical Version from Gandhāra of the ‘Miracle at Śrāvastī’ (Texts from the Split Collection 4). A Metrical Version from Gandhāra of the ‘Miracle at Śrāvastī’ (Texts from the Split Collection 4).” *Sōka daigaku kokusai bukkyōgaku kōtō kenkyūjo nenpō* 創価大学国際仏教学高等研究所年報 23: 3–42
- . 2022. “Avadāna Episodes (Texts from the Split Collection 5) Avadāna Episodes (Texts from the Split Collection 5).” *Sōka daigaku kokusai bukkyōgaku kōtō kenkyūjo nenpō* 創価大学国際仏教学高等研究所年報 25: 21–60
- Falk, Harry and Ingo Strauch. 2014. “The Bajaur and Split Collections of Kharoṣṭhī Manuscripts within the Context of Buddhist Gāndhārī Literature.” In Paul Harrison and Jens-Uwe Hartman, eds. *From Birch Bark to Digital Data: Recent Advances in Buddhist Manuscript Research. Papers Presented at the Conference Indic Buddhist Manuscripts: The State of the Field, Stanford, June 15-19 2009*. pp. 51-78. Wien: Österreichische Akademie der Wissenschaften
- Fallon, Paul D. 2002. *The Synchronic and Diachronic Phonology of Ejectives*. London and New York: Routledge.
- Fortson, Benjamin W. IV. 2010. *Indo-European Language and Culture: An Introduction* (2nd ed.). United Kingdom: Wiley-Blackwell.

- Fromkin, Victoria, Robert Rodman, Nina Hyams, Peter Collins, Mengistu Amberber, Felicity Cox and Rosalind Thornton. 2021. *An Introduction to Language* (10th ed.). Melbourne, Cengage Learning Australia.
- Fussman, Gérard. 1989. “Gāndhārī écrite, Gāndhārī parlée.” In Collette Caillat, ed., *Dialectes dans les littératures indo-aryennes*. Publications de l’Institut de civilisation indienne, série in-8°, fascicule 55, pp. 433-501. Paris: Institut de civilisation indienne.
- Geiger, Wilhelm. 1943. *Pāli Literature and Language*. Authorised English translation by Batakrishna Ghosh. Calcutta: Calcutta University Press.
- Glass, Andrew. 2000. *A Preliminary Study of Kharoṣṭhī Manuscript Paleography*. Masters Thesis. Department of Asian languages and Literature, The University of Washington.
- . 2007. *Four Gāndhārī Saṃyuktāgama Sūtras. Senior Kharoṣṭhī Fragment 5. Gandhāran Buddhist Texts 4*. Seattle and London: University of Washington Press.
- Gogoi, Poppy, Stephen Morey and Pittayawat Pittayaporn. 2020. “The Tai Ahom sound system as reflected by the texts recorded in the bark manuscripts.” *Journal of Southeast Asian Linguistics Society, JSEALS 13.2 (2020)*: pp. 14-42
- Gordon, Matthew K. 2016. *Phonological Typology*. Oxford: Oxford University Press.
- Grierson, G.A. 1913. “The Phonetics of the Wardak Vase.” *The Journal of the Royal Asiatic Society of Great Britain and Ireland*. Jan. 1913, pp. 141-144.
- . 1929. *Torwali. An account of a Dardic Language of the Swat Kohistan*. London: Royal Asiatic Society.
- Harrison, Paul, Timothy Lenz and Richard Salomon. 2018. “Fragments of a Gāndhārī Manuscript of the Pratyutpannabuddhasaṃmukhāvasthitasamādhisūtra (Studies in Gāndhārī Manuscripts 1).” *Journal of the International Association of Buddhist Studies* 41: 117–43.
- Hayes, Bruce. 1995. *Metrical Stress Theory. Principles and Case Studies*. Chicago and London: The University of Chicago Press
- Henke, Eric, Ellen M. Kaisse and Richard Wright. 2012. “Is the Sonority Sequencing Principle an epiphenomenon?” In Steve Parker ed., *The Sonority Controversy*. Berlin / Boston: De Gruyter Mouton.

- Hock, H. H. 2015. "The northwest of South Asia and beyond: The issue of Indo-Aryan retroflexion yet again." *Journal of South Asian Languages and Linguistics* 2(1) (2015) pp. 111-135.
- Houben, Jan E.M. 2018. "Linguistic Paradox and Diglossia: the emergence of Sanskrit and Sanskritic language in Ancient India. *Open Linguistics* 2018; 4: 1-18. De Gruyter. Retrieved from:
<https://www.degruyter.com/downloadpdf/j/opli.2018.4.issue-1/opli-2018-0001/opli-2018-0001.pdf>
- Hultzsch, E, ed. 1925. *Inscriptions of Aśoka*. Corpus Inscriptionum Indicarum Vol. 1. Printed for the Government of India. Oxford: Clarendon Press.
- International Phonetic Association. 1999. *Handbook of The International Phonetic Association: A guide to the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- International Phonetic Association. 2015. *The International Phonetic Alphabet (revised to 2015)*. Retrieved from International Phonetic Association website:
https://www.internationalphoneticassociation.org/sites/default/files/IPA_Kiel_2015.pdf
- Jain, Dhanesh. 2003. "Sociolinguistics of the Indo-Aryan languages." In Cardona, George, and Danesh Jain, eds. *The Indo-Aryan Languages*. pp. 46-66. London and New York: Routledge.
- Janson, Tore. 2012. *The History of Languages. An Introduction*. Oxford, Oxford University Press.
- Jantrasrisalai, Chanida, Timothy Lenz, Lin Qian and Richard Salomon. 2016. "Fragments of an Ekottarikāgama Manuscript in Gāndhārī." In Jens Braarvig, ed., *Buddhist Manuscripts, Volume IV. Manuscripts in the Schøyen Collection*, pp. 1–122. Oslo: Hermes Publishing.
- Jun, Jongho. 1995. *Perceptual and Articulatory Factors in Place Assimilation: An Optimality Theoretic Approach*. (PhD Dissertation). UCLA occasional Papers in Linguistics 16. Dissertation Series 2, December 1995. Los Angeles: University of California.
- Jun, Jongho. 1996. "Place assimilation is not the result of gestural overlap: evidence from Korean and English." *Phonology*, 13, pp. 377-407. Cambridge University Press.

- Jun, Jongho. 2004. "Place Assimilation." In Hayes, Bruce, Kirchner, Robert. & Steriade, Donca, eds. *Phonetically based phonology*. pp. 58-86. Cambridge, New York: Cambridge University Press.
- Karpik, Stefan. 2019a. "The Buddha taught in Pali: A working hypothesis". *Journal of the Oxford Centre for Buddhist Studies*, Vol. 17. November 2019. pp. 10-86.
- . 2019b. "A Reply to Bryan Levman's *The Language the Buddha Spoke*". *Journal of the Oxford Centre for Buddhist Studies*, Vol. 17. November 2019. pp. 106-116.
- Katre, Sumitra Mangesh. 1964. *Prakrit Languages and their contribution to Indian culture* (2nd ed.). Poona: Deccan College, Post-graduate and Research Institute.
- Katz, Jonah. 2016. "Lenition, perception and neutralization." *Phonology*, 33: pp. 43-85. Cambridge University Press.
- Kent, Ray D. and Charles Read. 2002. *The Acoustic Analysis of Speech* (2nd ed.). Canada: Delmar, Thomson Learning.
- Kirchner, Robert Martin. 1998. *An Effort-Based Approach to Consonant Lenition*. (PhD Dissertation). Department of Asian Languages and Literature, University of California, Los Angeles.
- . 2004. "Consonant Lenition" In Hayes, Bruce, Robert Kirchner and Donca Steriade, eds. *Phonetically based phonology*. pp. 313-345. Cambridge, New York: Cambridge University Press.
- Kobayashi, Masato. 2004. *Historical Phonology of Old Indo-Aryan Consonants*. Study of Languages and Cultures of Asia and Africa Monograph Series 42. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Konow, Sten, ed. 1929. *Kharoṣṭhī Inscriptions with the Exception of Those of Aśoka*. Corpus Inscriptionum Indicarum 2.1. Calcutta: Government of India. Reprint Varanasi: Indological Book House, 1969.
- Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford, UK and Cambridge, Massachusetts: Blackwell Publishers
- Lahiri, Aditi. 2015. Blog: Observing language change and language processing: old manuscripts, new brains. The British Academy.
<https://www.thebritishacademy.ac.uk/blog/observing-language-change-and-language-processing-old-manuscripts-new-brains/>

- Lee, Mei-huang. 2009. *A Study of the Gāndhārī Dārukkhandhopamasutta (Discourse on the Simile of the Log)*. PhD Dissertation. Department of Asian Languages and Literature, University of Washington.
- Lenz, Timothy. 2003. *A New Version of the Gāndhārī Dharmapada and a Collection of Previous–Birth Stories. British Library Kharoṣṭhī Fragments 16 + 25. Gandhāran Buddhist Texts 3*. Seattle and London: University of Washington Press.
- . 2010. *Gandhāran Avadānas. British Library Kharoṣṭhī Fragments 1-3 and 21 and Supplementary Fragments A-C. Gandhāran Buddhist Texts 6*. Seattle and London: University of Washington Press.
- Levman, Bryan. 2014. “Linguistic Ambiguities, the Transmissional Process, and the Earliest Recoverable Language of Buddhism.” *Dissertation, Department for the Study of Religion*, University of Toronto
- . 2016. “The Language of Early Buddhism.” *Journal of South Asian Languages and Linguistics* 3.1: 1–41.
- . 2019. “The Language the Buddha Spoke”. *Journal of the Oxford Centre for Buddhist Studies*, Vol. 17. November 2019. pp. 63-105.
- Lunsford, Wayne A. 2001. *An Overview of Linguistic Structures in Torwali, a Language of Northern Pakistan*. (Masters Thesis). The University of Texas at Arlington.
- Macdonell, A. A. 1910. *Vedic Grammar*. Strassburg. Verlag Von Karl J. Trübner.
- Macdonell, A. A. 1916. *A Vedic Grammar for Students*. London: Oxford University Press.
- Macdonell, Arthur A. 1927. *A Sanskrit Grammar for Students* (3rd ed.). London: Oxford University Press.
- Marino, Joseph. 2017. *Metaphor and Pedagogy in Early Buddhist Literature: An Edition and Study of Two Sūtras from the Senior Collection of Gāndhārī Manuscripts*. PhD Dissertation. Department of Asian Languages and Literature, University of Washington.
- . 2021a. “Mahaparaḍaha Sutra.” *Journal of Gandhāran Buddhist Texts*, March 20, 2021, <https://gandhari-texts.sydney.edu.au/edition/mahapara%e1%b8%8daha-sutra/>.
- . 2021b. “Suhadukha Sutra.” *Journal of Gandhāran Buddhist Texts*, March 20, 2021, <https://gandhari-texts.sydney.edu.au/edition/suhadukha-sutra/>.

- Masica, Colin P. 1991. *The Indo-Aryan Languages*. (1st paperback ed. 1993).
Cambridge: Cambridge University Press.
- Matsumoto, Kazuko. 2019. "Language variation and change." In Heinrich, Patrick and Yumiko Ohara (eds.). *Routledge handbook of Japanese sociolinguistics*. pp. 199-217. London: Routledge
- McCarthy, John J. 2008. "The gradual path to cluster simplification." *Phonology*, 25, pp 271-319. Cambridge University Press.
- McCrabb, Ian. 2021. *Buddha Bodies and the Benefits of Relic Establishment: Insights from a Digital Framework for the Analysis of Formulaic Sequences in Gāndhārī Relic Inscriptions*. PhD Dissertation. School of Languages and Culture, Faculty of Arts and Social Sciences, University of Sydney
- McWhorter, John. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford: Oxford University Press.
- Mehendale, Madhukar Anant. 1948. *Historical Grammar of Inscriptional Prakrits*. Pune: Akshar Sanchay. Reprint Poona: Deccan College Dissertation Series: 3. Deccan College Postgraduate and Research Institute, 1997.
- Michelson, Truman. 1909. "Linguistic Notes on the Shāhbāzgarhi and Mansehra Redactions of Asoka's Fourteen-Edicts." *The American Journal of Philology* 30: 416–29.
- Misra, Bal Govind. 1967. *Historical Phonology of the Modern Standard Hindi: Proto-Indo-European to the Present*. (PhD Dissertation). Faculty of the Graduate School, Cornell University. Retrieved from: <https://search-proquest-com.ezproxy1.library.usyd.edu.au/docview/302238511?pq-origsite=summon>
- Monier-Williams, M. 1899. *A Sanskrit-English Dictionary*. Oxford: Clarendon Press
- Murray, Robert W. 1982. "Consonant Cluster Developments in Pāli." *Folia Linguistica Historica* III/2, pp 163-184.
- Nettle, Daniel and Suzanne Romaine. 2000. *Vanishing voices. The Extinction of the World's Languages*. (1st paperback ed. 2002). Oxford, Oxford University Press.
- Norman, K.R. 1992. "The origin of Pāli and its position among the Indo-European languages." In *Journal of Pali and Buddhist Studies* I, pp. 1-27. Rpt. In Norman K.R. 1992. *Collected Papers, Volume III*, pp 225-243. Oxford: the Pali Text Society.

- . 1993. “The languages of early Buddhism.” In *Premier Colloque Etienne Lamotte*. (Louvain 1993, pp. 83-89) (Publications de l’Institute Orientaliste de Louvain. Rpt. In Norman K.R. 1994. *Collected Papers, Volume V*. pp 146-168. Oxford: the Pali Text Society.
- . 1994. “The Aśokan Inscriptions and Prakrit dialect geography.” In N. N. Bhattacharyya (ed.): *Jainism and Prakrit in Ancient and Medieval India: Essays for Prof. Jagdish Jain*, New Delhi 1994, pp. 51-57. Rpt. In Norman K.R. 1996. *Collected Papers, Volume VI*, pp. 1-8. Oxford: the Pali Text Society.
- . 2006. “Buddhism and regional dialects.” In *A Philological Approach to Buddhism. The Bukkyō Dendō Kyōkai Lectures 1994*, (2nd ed.) Ch. IV, pp. 75-97. Oxford: the Pali Text Society.
- Oberlies, Thomas. 1999. “Middle Indo-Aryan and (the) Vedic (Dialects) (Miscellanea Pāli VII)” *Historische Sprachforschung (Historical Linguistics)* 112, pp. 39-57
- . 2003a. *A Grammar of Epic Sanskrit*. Wezler, Albrecht, and Michael Witzel (eds.) Indian Philology and South Asian Studies, Vol. 5. Berlin and New York: Walter de Gruyter
- . 2003b. “Aśokan Prakrit and Pāli.” In Jain, Danesh, and George Cardona, eds. *The Indo-Aryan Languages*. pp. 161-203. London and New York: Routledge.
- . 2019. *Pāli Grammar. The Language of the canonical texts of Theravāda Buddhism, Volumes I and II*. Bristol: The Pali Text Society.
- Oh, Grace E., and Melissa A Redford. 2012. “The production and phonetic representation of fake geminates in English.” *Journal of Phonetics*, 40 (2012). pp.82-91
- Ollett, Andrew. 2017. *Language of the Snakes: Prakrit, Sanskrit, and the Language Order of Premodern India*. Oakland California: University of California Press.
- Padgett, Jaye. 1994. “Stricture and Nasal Place Assimilation.” *Natural Language and Linguistic Theory*, 12, pp 465-513. The Netherlands: Kluwer Academic Publishers.
- Pischel, R. 1879. *Comparative Grammar of the Prākṛit Languages*. Translated from the German by Subhadra Jhā. Reprint Varanasi, Delhi and Patna: Motilal Banarasi Das, 1957.
- Rhys Davids, T. W. ed. 1921-1925. *The Pali Text Society’s Pali-English Dictionary*. London. The Pali Text Society.

- Salomon, Richard. 1988. "Two New Kharoṣṭhī documents from Central Asia" *Central Asiatic Journal* 32: 98-108.
- . 1995. "On the Origin of the Early Indian Scripts" *Journal of the American Oriental Society*, 115.2: 271-279.
- . 1998a. "Kharoṣṭhī Manuscript Fragments in the Pelliot Collection, Bibliothèque Nationale de France" *Bulletin d'études indiennes* 16: 123-160.
- . 1998b. *Indian Epigraphy. A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and the Other Indo-Aryan Languages*. Oxford and New York: Oxford University Press.
- . 1999. *Ancient Buddhist Scrolls from Gandhāra. The British Library Kharoṣṭhī Fragments*. Seattle: The University of Washington Press
- . 2000. *A Gāndhārī Version of the Rhinoceros Sutra. British Library Kharoṣṭhī Fragment 5B. Gandhāran Buddhist Texts 1*. Seattle and London: University of Washington Press.
- . 2001. "Gāndhārī Hybrid Sanskrit": New Sources for the Study of the Sanskritisation of Buddhist Literature. *Indo-Iranian Journal* 44 (2001). pp 241-252.
- . 2002. "Gāndhārī and the other Indo-Aryan languages in the light of newly-discovered Kharoṣṭhī manuscripts." *Proceedings of the British Academy*, 116: 119-134.
- . 2003a. "The Senior Manuscripts: Another Collection of Gandhāran Buddhist Scrolls." *Journal of the American Oriental Society*, 123.1: 73-92.
- . 2003b. "Writing Systems of the Indo-Aryan Languages" In Cardona, George, and Danesh Jain, eds. *The Indo-Aryan Languages*. pp. 67-103. London and New York: Routledge.
- . 2008. *Two Gāndhārī Manuscripts of the Songs of Lake Anavatapta (Anavatapta-gāthā). British Library Kharoṣṭhī Fragment 1 and Senior Scroll 14. Gandhāran Buddhist Texts 5*. Seattle and London: University of Washington Press.
- . 2014. "Gāndhārī Manuscripts in the British Library, Schøyen and Other Collections." In Paul Harrison and Jens-Uwe Hartman, eds. *From Birch Bark to Digital Data: Recent Advances in Buddhist Manuscript Research. Papers Presented at the Conference Indic Buddhist Manuscripts: The State of the Field*,

- Stanford, June 15-19 2009. pp. 1-18. Wien: Österreichische Akademie der Wissenschaften.
- . 2018. *The Buddhist Literature of Ancient Gandhāra: An Introduction with Selected Translations*. Classics of Indian Buddhism. Somerville: Wisdom Publications.
- Salomon, Richard, and Stefan Baums. 2007. “Sanskrit Ikṣvaku, Pali Okkāka, and Gāndhārī Iṣmaho.” *Journal of the Pali Text Society*, 29: pp. 201-227.
- Schlosser, Andrea. 2016. *On the Bodhisattva Path in Gandhāra: Edition of Fragment 4 and 11 from the Bajaur Collection of Kharoṣṭhī Manuscripts*. PhD Dissertation. Zur Erlangung des Doktorgrades eingereicht am Fachbereich Geschichts- und Kulturwissenschaften der Freien Universität Berlin im Dezember 2013, Überarbeitete Fassung: Februar 2106.
- . 2022. *Three Early Mahāyāna Treatises from Gandhāra: Bajaur Kharoṣṭhī Fragments 4, 6, 11. Gandhāran Buddhist Texts 7*. Seattle: University of Washington Press.
- Schlosser, Andrea and Ingo Strauch. 2016a. “Abhidharmic Elements in Gandhāran Mahāyāna Buddhism: Groups of Four and the *abhedyaprasādas* in the Bajaur Mahāyāna Sūtra.” In: Bart Dessein and Weijen Teng, eds., *Text, History, and Philosophy: Abhidharma across Buddhist Scholastic Traditions*. Brill’s Indological Library, Volume 50, pp. 47–107. Leiden: Brill.
- . 2016b. “The Bajaur Mahāyāna Sūtra. A preliminary analysis of its contents.” *Journal of the International Association of Buddhist Studies*, Volume 39 2016, pp 309-335.
- Schmeiser, Benjamin. 2008. “A Gestural Approach to Gemination and Nasals in Pali.” *Kansas Working Papers in Linguistics*, Vol. 30 (2008), pp. 302-314.
- Schneider, Edgar W. 2002. “Investigating Variation and Change in Written Documents.” In Chambers, J. K., Peter Trudgill and Natalie Schilling (eds.), *The Handbook of Language Variation and Change*, pp. 67-96. Massachusetts, Oxford. Blackwell.
- Silverlock, Blair Alan. 2015. *An Edition and Study of the Goṣiga-sutra, the Cow-Horn Discourse (Senior Collection scroll no 12): An Account of the Harmonious Aṇarudha Monks*. PhD Dissertation. Department of Indian Subcontinent Studies, University of Sydney

- Strauch, Ingo. 2008a. *The Bajaur Collection: A new collection of Kharoṣṭhī manuscripts – A preliminary catalogue and survey*-. Online version 1.1 (May 2008). Retrieved from:
https://www.academia.edu/25779968/The_Bajaur_collection_A_new_collection_of_Kharoṣṭhī_manuscripts_A_preliminary_catalogue_and_survey
- . 2008b. “The Bajaur Collection of Kharoṣṭhī manuscripts – a preliminary survey.” *Studien zur Indologie und Iranistik* 25 (2008) pp 103-136.
- . 2012a. “Catalogue of Inscriptions and Drawings in the Cave Hoq.” In: Ingo Strauch, ed., *Foreign Sailors on Socotra: The Inscriptions and Drawings from the Cave Hoq*. Vergleichende Studien zu Antike und Orient, Band 3, pp. 25–218. Bremen: Hempen Verlag.
- . 2012b. “The character of the Indian Kharoṣṭhī script and the “Sanskrit Revolution”: a writing system between identity and assimilation.” In Alex de Voogt and Joachim Friedrich Quack eds. *The Idea of Writing: Writing Across Borders*. pp. 131-168. Leiden, Boston: Brill
- . 2014. “The Bajaur Collection of Kharoṣṭhī Manuscripts: Mahāprajāpatī Gautamī and the order of nuns in a Gandhāran version of the Dakṣiṇāvibhaṅgasūtra.” In Alice Collett, ed. *Women in Early Indian Buddhism. Comparative Textual Studies*. pp 17- 45, 247-266. New York: Oxford University Press
- Suzuki, Yasuko. 2002a. “Consonant Cluster Changes in Pali: Toward Restricting the Phonological Patterns (Part 1).” *Journal of Inquiry and Research*, Vol 75, pp. 97-125. Kansai GaiDai University.
- . 2002b. “Consonant Cluster Changes in Pali: Toward Restricting the Phonological Patterns (Part 2).” *Journal of Inquiry and Research*, Vol 76, pp. 63-86. Kansai GaiDai University.
- . 2012. “Sanskrit Geminatio as Lengthening: In Relation to Svarabhakti, Yama, and Middle Indo-Aryan Assimilation.” In Jared S. Klein and Kazuhiko Yoshida eds., *Indic across the Millenia: from the Rigveda to Modern Indo-Aryan: 14th World Sanskrit Conference, Kyoto, Japan, September 1st-5th, 2009*. pp.193-208 Bremen: Hempen Verlag.
- . 2015. “On characterizing Sanskrit Geminatio.” *Journal of Inquiry and Research*, Vol 102, pp. 1-18. Kansai GaiDai University.

- . 2019. “The Development of y-final clusters in Aśokan Rock Edicts.” *Journal of Inquiry and Research*, Vol 110, pp. 19-34. Kansai GaiDai University.
- . 2020. “The Development of Stop Clusters in Aśokan Rock Edicts.” *Journal of Inquiry and Research*, Vol 111, pp. 23-37. Kansai GaiDai University.
- . 2021. “The Development of Labial Clusters in Aśokan Rock Edicts.” *Studio Orientalia Electronica*, 9(1) (2021) pp.160-172.
- Trudgill, Peter. 2002. “Linguistic and Social Typology.” In Chambers, J. K., Peter Trudgill, and Natalie Schilling-Estes (eds.). *The Handbook of Language Variation and Change*. pp 707-728.
- . 2011. *Sociolinguistic typology: social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. “Empirical Foundations for a Theory of Language Change” In Lehman, Winifred P., and Yakov Malkiel (eds.), *Directions for Historical Linguistics*. pp 95-195. Austin: University of Texas Press.
- Whitney, William Dwight. 1924. *Sanskrit Grammar. Including both the Classical Language and the Older Dialects of Veda and the Brahmana*. 5th Ed. London. Turner & Co.
- Witzel, Michael. 1989. “Tracing the Vedic Dialects.” In C. Caillat, ed. *Dialectes dans les littératures indo-aryennes*. pp. 97-264. Paris: P., Collège de France - Diff. De Boccard.
- . 1999. “Substrate Languages in Old Indo-Aryan (Ṛgvedic, Middle and Late Vedic)” *Electronic Journal of Vedic Studies (EJVS)* 5-1 (1999) pp 1-67.
- Woolner, Alfred C. 1917. *Introduction to Prakrit*. Lahore: University of the Panjab.
- . 1924. *Asoka Text and Glossary*. Punjab University Oriental Publications. Calcutta. Humphrey Milford, Oxford University Press.
- Wright, Richard. 2004. “A review of perceptual cues and cue robustness.” In Hayes, Bruce, Robert Kirchner and Donca Steriade, eds. *Phonetically based phonology*. pp. 34-57. Cambridge, New York: Cambridge University Press.
- Wynne, Alexander. 2019. “Editorial. Once more on the language of the Buddha”. *Journal of the Oxford Centre for Buddhist Studies*, Vol. 17. November 2019. pp. 8-10.

Zwalf, W. 1996. *A catalogue of the Gandhāra Sculpture in the British Museum.*
Volume I: Text. London: British Museum Press.